Austrian Journal of Statistics
2024, Volume 53, 120–134.
http://www.ajs.or.at/
doi:10.17713/ajs.v53i1.1635

Bayes Estimation for the Mean Matrix of the SSMESN Family of Matrix Variate Distributions

Amir Rezaei
Ph.D. graduate in Statistics
from the University of Birjand
Birjand, Iran

Fatemeh Yousefzadeh
Department of Statistics
University of Birjand
Birjand, Iran

Abstract

In this paper, the problem of finding a Bayes estimation for the mean matrix of the scale and shape mixtures of matrix variate extended skew normal distributions is considered, and its applications in the multivariate linear regression and the stress-strength models are described. Finally, a simulation study and a real data analysis are presented for applications.

Keywords: matrix variate distributions, Bayes estimation, multivariate linear regression model, stress-strength model.

1. Introduction

The matrix variate distributions have a very important role in multivariate analysis methods; for example, the distribution of the maximum likelihood estimator of the covariance matrix of a multivariate normal distribution is the Wishart distribution, which plays a pivotal role in related analysis. The matrix variate normal distribution is one of the most important matrix variate distributions; for more about this distribution, see Gupta and Nagar (1999) and Gupta, Varga, and Bodnar (2013). A $p \times n$ random matrix $\boldsymbol{X}$ follows a matrix variate normal distribution if its probability density function (PDF) can be written as

$$\phi_{p\times n}(\boldsymbol{X}; \boldsymbol{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}) = (2\pi)^{-\frac{pn}{2}} |\boldsymbol{\Psi}|^{-\frac{p}{2}} |\boldsymbol{\Sigma}|^{-\frac{n}{2}} etr\Big\{ -\frac{1}{2}\boldsymbol{\Psi}^{-1}(\boldsymbol{X}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{M})\Big\},$$

where $etr\{A\} = \exp\{\text{tr}(A)\}$, $\boldsymbol{M}$ is a $p \times n$ mean matrix, $\boldsymbol{\Sigma}$ is a $p \times p$ positive definite matrix and $\boldsymbol{\Psi}$ is an $n \times n$ positive definite matrix. The normal matrix variate $\boldsymbol{X}$ is denoted by $\boldsymbol{X} \sim N_{p\times n}(\boldsymbol{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$.

There are different skew versions of the matrix variate normal distribution. One of these skew versions is the matrix variate extended skew normal distribution, which was introduced by Ning and Gupta (2012). A $p \times n$ random matrix $\boldsymbol{X}$ is said to follow a matrix variate extended skew normal distribution with a $p \times n$ mean matrix $\boldsymbol{M}$, a $p \times p$ positive definite matrix $\boldsymbol{\Sigma}$

and $n \times n$ positive definite matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Psi}$, if its PDF is

$$f_{ESN}(\boldsymbol{X};\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma},\boldsymbol{\Omega},\boldsymbol{\lambda},\boldsymbol{\delta}) = \frac{\phi_{p\times n}(\boldsymbol{X};\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})}{\Phi_n(\boldsymbol{\delta};\boldsymbol{\Omega}+\boldsymbol{\lambda}'\boldsymbol{\lambda}\boldsymbol{\Psi})}\Phi_n(\boldsymbol{\delta}+(\boldsymbol{X}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega}),$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\delta}$ are p and q dimensional vectors, respectively, $\phi_{p\times n}(\cdot;\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})$ is the PDF of the matrix variate normal distribution $N_{p\times n}(\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})$ and $\Phi_n(\cdot;\boldsymbol{\Omega})$ is the cumulative distribution function (CDF) of the multivariate normal distribution $N_n(\boldsymbol{0},\boldsymbol{\Omega})$. The extended skew normal matrix variate $\boldsymbol{X}$ is denoted by $\boldsymbol{X}\sim ESN_{p\times n}(\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma},\boldsymbol{\Omega},\boldsymbol{\lambda},\boldsymbol{\delta})$.

Recently the scale and shape mixtures of matrix variate extended skew normal (SSMESN) distributions was introduced by Rezaei, Yousefzadeh, and Arellano-Valle (2020) as a new family of matrix variate distributions which includes a wide range of distributions such as matrix variate normal, matrix variate skew normal, matrix variate t, matrix variate skew t, matrix variate skew-t-normal and matrix variate skew-normal-Cauchy distributions. A $p\times n$ random matrix $\boldsymbol{Y}$ follows an SSMESN distribution with a $p\times n$ mean matrix $\boldsymbol{M}$, a $p\times p$ positive definite matrix $\boldsymbol{\Sigma}$ and $n\times n$ positive definite matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Psi}$ if

$$\boldsymbol{Y}\mid\theta=\theta_0,\omega=\omega_0\sim ESN_{p\times n}(\boldsymbol{M},\boldsymbol{\Psi}\otimes k(\theta_0)\boldsymbol{\Sigma},\boldsymbol{\Omega},s(\theta_0,\omega_0)\boldsymbol{\lambda},\boldsymbol{\delta}),$$

or equivalently, if its PDF is as follows

$$f(\boldsymbol{Y};\boldsymbol{M},\boldsymbol{\Sigma},\boldsymbol{\Psi},\boldsymbol{\Omega},\boldsymbol{\lambda},\boldsymbol{\delta}) = \int_{S_Q} f_{ESN}(\boldsymbol{Y};\boldsymbol{M},\boldsymbol{\Psi}\otimes k(\theta)\boldsymbol{\Sigma},\boldsymbol{\Omega},s(\theta,\omega)\boldsymbol{\lambda},\boldsymbol{\delta})dQ(\theta,\omega),$$

where θ and ω are two random variables that have joint distribution $Q(\theta,\omega)$ with support S_Q and marginal distributions $H(\theta)$ and $G(\omega)$, $k(\theta)$ is a weight function and $s(\theta,\omega)$ is a real-valued function. The SSMESN matrix variate $\boldsymbol{Y}$ is denoted by $\boldsymbol{Y}\sim SSMESN_{p\times n}(\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma},\boldsymbol{\Omega},\boldsymbol{\lambda},\boldsymbol{\delta};(k,s),Q)$.

When $\boldsymbol{M}=\boldsymbol{1}_n'\otimes\boldsymbol{\mu}$, $\boldsymbol{\delta}=\delta\boldsymbol{1}_n$ and $\boldsymbol{\Omega}=\boldsymbol{\Psi}=\boldsymbol{I}_n$, where $\boldsymbol{\mu}\in\mathbb{R}^p$, $\delta\in\mathbb{R}^1$ and $\boldsymbol{1}_n$ is an n-dimensional vector of ones, an important situation occurs for the SSMESN matrix variate $\boldsymbol{Y}$ with the columns $\boldsymbol{y}_1,\dots,\boldsymbol{y}_n$. In this situation,

$$\boldsymbol{y}_i\mid\theta=\theta_0,\omega=\omega_0\overset{iid}{\sim}ESN_p(\boldsymbol{\mu},k(\theta_0)\boldsymbol{\Sigma},s(\theta_0,\omega_0)\boldsymbol{\lambda},\delta),\quad i=1,\dots,n,$$

with the conditional PDF

$$\begin{aligned}&f_{ESN}(\boldsymbol{y}_i\mid\theta_0,\omega_0;\boldsymbol{\mu},k(\theta_0)\boldsymbol{\Sigma},s(\theta_0,\omega_0)\boldsymbol{\lambda},\delta)\\&=\frac{1}{\Phi_1(\delta/\sqrt{1+s(\theta_0,\omega_0)^2\boldsymbol{\lambda}'\boldsymbol{\lambda}})}\phi_p(\boldsymbol{y}_i;\boldsymbol{\mu},k(\theta_0)\boldsymbol{\Sigma})\\&\times\Phi_1(\delta+s(\theta_0,\omega_0)k(\theta_0)^{-\frac{1}{2}}(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}),\quad\boldsymbol{y}_i\in\mathbb{R}^p,\end{aligned}$$

where ϕ_p and Φ_1 are the PDF of the p-variate normal distribution and the CDF of the univariate standard normal distribution, respectively.

The matrix variate SSMESN family includes some different matrix variate distributions, and is a quite large family of this type of distributions. For example,

- If $k(\theta)=s(\theta,\omega)=1$ and $\boldsymbol{\lambda}=\boldsymbol{0}$, then we have the matrix variate normal distribution.
- If $\boldsymbol{\lambda}=\boldsymbol{0}$, then we obtain the scale mixture of matrix variate normal distributions which proposed by Gupta and Varga (1995). We denote this subfamily by $SMN_{p\times n}(\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma};k,H)$.
- If $k(\theta)=s(\theta,\omega)=1$, the matrix variate extended skew normal distribution is obtained.
- If $\boldsymbol{\delta}=\boldsymbol{0}$, then the SSMESN matrix variate $\boldsymbol{Y}$ follows the matrix variate skew t distribution with ν degrees of freedom by considering $k(\theta)=\theta$ and $s(\theta,\omega)=1$ with $\theta\sim IGamma(\frac{\nu}{2},\frac{\nu}{2})$, where $IGamma(a,b)$ denotes the inverse gamma distribution with shape parameter a and scale parameter b. We use the notation $ST_{p\times n}(\boldsymbol{M},\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma},\boldsymbol{\Omega},\boldsymbol{\lambda},\nu)$ to denote this distribution.

- If $\boldsymbol{\delta} = \mathbf{0}$, $\boldsymbol{\Psi} = \boldsymbol{\Omega} = \boldsymbol{I}_n$, $k(\theta) = 1$ and $s(\theta, \omega) = \omega^{-\frac{1}{2}}$ with $\omega \sim IGamma(\frac{1}{2}, \frac{1}{2})$, then the SSMESN matrix variate $\boldsymbol{Y}$ follows the matrix variate skew-normal-Cauchy distribution which is denoted here by $SNC_{p\times n}(\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$.

In the next section, a posterior density for the mean matrix of the matrix variate SSMESN distributions is obtained and some of its particular cases are provided. In Section 3, applications of the obtained results in the multivariate linear regression and the stress-strength models will be discussed. Sections 4 and 5 will present a simulation study for comparing the Bayes estimations of a stress-strength reliability and a real data analysis for a multivariate linear regression model, respectively.

2. Posterior densities

It must minimize the posterior risk to find a Bayes estimator for a parameter. Therefore, in the first step, related posterior distribution or related posterior density should be obtained. In this section, by considering a matrix variate normal distribution as prior for the mean matrix of the matrix variate SSMESN distributions, a posterior density is derived for it. The result is given in the following proposition.

Proposition 2.1. *Suppose that $\boldsymbol{Y} \sim SSMESN_{p\times n}(\boldsymbol{M}, \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\lambda}, \boldsymbol{\delta}; (k, s), Q)$ where $\boldsymbol{\Sigma}$, $\boldsymbol{\Psi}$, $\boldsymbol{\Omega}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\delta}$ are known. If $\boldsymbol{M}$ is independent of θ and ω, and has prior distribution as $N_{p\times n}(\mathbf{0}_{p\times n}, \boldsymbol{\Psi}\otimes\boldsymbol{\Delta})$, where $\boldsymbol{\Delta}_{p\times p}$ is a positive definite matrix, then the posterior density of $\boldsymbol{M}$ is*

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto \int_{S_Q} \varrho_\theta |\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}} \frac{\phi_{p\times n}(\boldsymbol{M}; \boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}, \boldsymbol{\Psi}\otimes k(\theta)\boldsymbol{\Lambda}_\theta)}{\Phi_n(\boldsymbol{\delta}; \boldsymbol{\Omega} + s(\theta,\omega)^2\boldsymbol{\lambda}'\boldsymbol{\lambda}\boldsymbol{\Psi})} \times \Phi_n(\boldsymbol{\delta} + s(\theta,\omega)k(\theta)^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{\Omega}) dQ(\theta,\omega), \tag{1}$$

where $\boldsymbol{\Lambda}_\theta = (\boldsymbol{\Sigma}^{-1} + k(\theta)\boldsymbol{\Delta}^{-1})^{-1}$, $\boldsymbol{\tau} = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Psi}^{-1}$ and $\varrho_\theta = etr\Big\{\frac{\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}\boldsymbol{\tau}' - \boldsymbol{\tau}\boldsymbol{Y}'}{2k(\theta)}\Big\}$.

Proof. Since $\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto f(\boldsymbol{Y}|\boldsymbol{M})\pi(\boldsymbol{M})$, by the PDFs of the random matrices $\boldsymbol{Y}$ and $\boldsymbol{M}$, we can write

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto \int_{S_Q} \frac{(2\pi)^{-\frac{np}{2}}|\boldsymbol{\Psi}|^{-\frac{p}{2}}k(\theta)^{-\frac{np}{2}}}{\Phi_n(\boldsymbol{\delta}; \boldsymbol{\Omega} + s(\theta,\omega)^2\boldsymbol{\lambda}'\boldsymbol{\lambda}\boldsymbol{\Psi})} \times etr\Big\{\frac{-1}{2k(\theta)}\boldsymbol{\Psi}^{-1}(\boldsymbol{M}-\boldsymbol{Y})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{M}-\boldsymbol{Y}) - \frac{1}{2}\boldsymbol{\Psi}^{-1}\boldsymbol{M}'\boldsymbol{\Delta}^{-1}\boldsymbol{M}\Big\} \times \Phi_n(\boldsymbol{\delta} + s(\theta,\omega)k(\theta)^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{\Omega}) dQ(\theta,\omega).$$

Now, since

$$etr\Big\{\frac{-1}{2k(\theta)}\boldsymbol{\Psi}^{-1}(\boldsymbol{M}-\boldsymbol{Y})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{M}-\boldsymbol{Y})\Big\} = etr\Big\{\frac{-1}{2k(\theta)}\boldsymbol{\tau}\boldsymbol{Y}'\Big\} \times etr\Big\{-\frac{1}{2k(\theta)}\boldsymbol{\Psi}^{-1}\boldsymbol{M}'\boldsymbol{\Sigma}^{-1}\boldsymbol{M} + \frac{1}{k(\theta)}\boldsymbol{\tau}'\boldsymbol{M}\Big\},$$

and

$$etr\Big\{\frac{-1}{2k(\theta)}[\boldsymbol{\Psi}^{-1}\boldsymbol{M}'\boldsymbol{\Lambda}_\theta^{-1}\boldsymbol{M} - 2\boldsymbol{\tau}'\boldsymbol{M}]\Big\} = etr\Big\{\frac{1}{2k(\theta)}\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}\boldsymbol{\tau}'\Big\} \times etr\Big\{\frac{-1}{2k(\theta)}\boldsymbol{\Psi}^{-1}(\boldsymbol{M}-\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}'\boldsymbol{\Psi})'\boldsymbol{\Lambda}_\theta^{-1}(\boldsymbol{M}-\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}'\boldsymbol{\Psi})\Big\},$$

we have

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto \int_{S_Q} etr\Big\{\frac{\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}\boldsymbol{\tau}' - \boldsymbol{\tau}\boldsymbol{Y}'}{2k(\theta)}\Big\}|\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}} \frac{\phi_{p\times n}(\boldsymbol{M}; \boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}, \boldsymbol{\Psi}\otimes k(\theta)\boldsymbol{\Lambda}_\theta)}{\Phi_n(\boldsymbol{\delta}; \boldsymbol{\Omega} + s(\theta,\omega)^2\boldsymbol{\lambda}'\boldsymbol{\lambda}\boldsymbol{\Psi})} \times \Phi_n(\boldsymbol{\delta} + s(\theta,\omega)k(\theta)^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{\Omega}) dQ(\theta,\omega).$$

□

Note. By substituting $\boldsymbol{U} = \boldsymbol{M} - \boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}$, and using Lemma 2.1 of Harrar and Gupta (2008), the explicit form of $\pi(\boldsymbol{M}|\boldsymbol{Y})$ can be obtained as follows

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) = \frac{\int_{S_Q} \frac{\varrho_\theta|\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}}\phi_{p\times n}(\boldsymbol{M};\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi},\boldsymbol{\Psi}\otimes k(\theta)\boldsymbol{\Lambda}_\theta)\Phi_n(\boldsymbol{\delta}+s(\theta,\omega)k(\theta)^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega})}{\Phi_n(\boldsymbol{\delta};\boldsymbol{\Omega}+s(\theta,\omega)^2\boldsymbol{\lambda}'\boldsymbol{\lambda}\boldsymbol{\Psi})} dQ(\theta,\omega)}{\int_{S_Q} \frac{\varrho_\theta|\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}}\Phi_n(\boldsymbol{\delta}+s(\theta,\omega)k(\theta)^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega}+s(\theta,\omega)^2\boldsymbol{\lambda}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Lambda}_\theta\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\boldsymbol{\Psi})}{\Phi_n(\boldsymbol{\delta};\boldsymbol{\Omega}+s(\theta,\omega)^2\boldsymbol{\lambda}'\boldsymbol{\lambda}\boldsymbol{\Psi})} dQ(\theta,\omega)}.$$

The following corollaries can be written by using Proposition 2.1.

Corollary 2.1. *Let* $\boldsymbol{\Lambda} = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Delta}^{-1})^{-1}$.

(i) If $\boldsymbol{Y} \sim N_{p\times n}(\boldsymbol{M}, \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})$, *then* $\boldsymbol{M}|\boldsymbol{Y} \sim N_{p\times n}(\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \boldsymbol{\Psi}\otimes\boldsymbol{\Lambda})$.

(ii) If $\boldsymbol{Y} \sim ESN_{p\times n}(\boldsymbol{M}, \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\lambda}, \boldsymbol{\delta})$, *then the posterior distribution of* $\boldsymbol{M}$ *is*

$$ESN_{p\times n}(\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \boldsymbol{\Psi}\otimes\boldsymbol{\Lambda}, \boldsymbol{\Omega}, -\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}, \boldsymbol{\delta} + (\boldsymbol{Y} - \boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}).$$

Proof. As the matrix variate normal distribution is a special case of the matrix variate extended skew normal distribution, we only prove *(ii)*. Consider the exact form of posterior density in Note 2. If $k(\theta) = s(\theta,\omega) = 1$, then

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) = \frac{\phi_{p\times n}(\boldsymbol{M};\boldsymbol{\Lambda}\boldsymbol{\tau}\boldsymbol{\Psi}, \boldsymbol{\Psi}\otimes\boldsymbol{\Lambda})\Phi_n(\boldsymbol{\delta} + (\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega})}{\Phi_n(\boldsymbol{\delta} + (\boldsymbol{Y}-\boldsymbol{\Lambda}\boldsymbol{\tau}\boldsymbol{\Psi})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega} + \boldsymbol{\lambda}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\boldsymbol{\Psi})},$$

where $\boldsymbol{\tau} = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Psi}^{-1}$. Since $\boldsymbol{\Lambda}\boldsymbol{\tau}\boldsymbol{\Psi} = \boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}$, $\boldsymbol{\lambda}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} = (-\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda})'(-\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda})$, and

$$\begin{aligned}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} &= \boldsymbol{Y}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} - \boldsymbol{M}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\\ &= \boldsymbol{Y}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} - \boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} + \boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} - \boldsymbol{M}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\\ &= (\boldsymbol{Y}' - \boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda})\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} - (\boldsymbol{M}' - \boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda})\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\\ &= (\boldsymbol{Y} - \boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} + (\boldsymbol{M} - \boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Lambda}^{-\frac{1}{2}}(-\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}),\end{aligned}$$

we have

$$\begin{aligned}\pi(\boldsymbol{M}|\boldsymbol{Y}) =& \frac{\phi_{p\times n}(\boldsymbol{M};\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \boldsymbol{\Psi}\otimes\boldsymbol{\Lambda})}{\Phi_n\Big(\boldsymbol{\delta} + (\boldsymbol{Y}-\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega} + (-\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda})'(-\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda})\boldsymbol{\Psi}\Big)}\\ &\times \Phi_n\Big(\boldsymbol{\delta} + (\boldsymbol{Y}-\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda} + (\boldsymbol{M}-\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Lambda}^{-\frac{1}{2}}(-\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda});\boldsymbol{\Omega}\Big)\\ =& f_{ESN}(\boldsymbol{M};\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \boldsymbol{\Psi}\otimes\boldsymbol{\Lambda}, \boldsymbol{\Omega}, -\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}, \boldsymbol{\delta} + (\boldsymbol{Y}-\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}).\end{aligned}$$

□

Corollary 2.2. *If* $\boldsymbol{Y} \sim ST_{p\times n}(\boldsymbol{M}, \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\lambda}, \nu)$, *then the posterior density of* $\boldsymbol{M}$ *is as follows:*

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto E_\theta\left[\varrho_\theta|\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}}\phi_{p\times n}(\boldsymbol{M};\boldsymbol{\Lambda}_\theta\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \boldsymbol{\Psi}\otimes\theta\boldsymbol{\Lambda}_\theta)\Phi_n(\theta^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega})\right],$$

where $\theta \sim IGamma(\frac{\nu}{2}, \frac{\nu}{2})$, $\boldsymbol{\Lambda}_\theta = (\boldsymbol{\Sigma}^{-1} + \theta\boldsymbol{\Delta}^{-1})^{-1}$ *and* $\varrho_\theta = etr\Big\{\frac{\boldsymbol{\Lambda}_\theta\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Psi}^{-1}\boldsymbol{Y}'\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Psi}^{-1}\boldsymbol{Y}'}{2\theta}\Big\}$.

Proof. Matrix variate skew t distribution is one of the distributions belonging to the matrix variate SSMESN family in which $\boldsymbol{\delta} = \boldsymbol{0}$, $s(\theta,\omega) = 1$ and $k(\theta) = \theta$ with $\theta \sim IGamma(\frac{\nu}{2}, \frac{\nu}{2})$. Hence, the posterior density in Proposition 2.1 becomes as follows

$$\begin{aligned}\pi(\boldsymbol{M}|\boldsymbol{Y}) &\propto \int_{S_H} \varrho_\theta|\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}}\phi_{p\times n}(\boldsymbol{M};\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}, \boldsymbol{\Psi}\otimes\theta\boldsymbol{\Lambda}_\theta)\Phi_n(\boldsymbol{\delta} + \theta^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega})dH(\theta)\\ &= E_\theta\left[\varrho_\theta|\boldsymbol{\Lambda}_\theta|^{\frac{n}{2}}\phi_{p\times n}(\boldsymbol{M};\boldsymbol{\Lambda}_\theta\boldsymbol{\tau}\boldsymbol{\Psi}, \boldsymbol{\Psi}\otimes\theta\boldsymbol{\Lambda}_\theta)\Phi_n(\boldsymbol{\delta} + \theta^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda};\boldsymbol{\Omega})\right],\end{aligned}$$

where H is the CDF of the distribution $IGamma(\frac{\nu}{2}, \frac{\nu}{2})$ with the support $S_H = [0, +\infty)$ and $\varrho_\theta = etr\left\{\frac{\boldsymbol{\Lambda}_\theta \boldsymbol{\tau}\boldsymbol{\Psi}\boldsymbol{\tau}' - \boldsymbol{\tau}\boldsymbol{Y}'}{2k(\theta)}\right\}$. The proof is completed by $\boldsymbol{\tau}\boldsymbol{\Psi} = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}$, $\boldsymbol{\tau}\boldsymbol{\Psi}\boldsymbol{\tau}' = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Psi}^{-1}\boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}$, and $\boldsymbol{\tau}\boldsymbol{Y}' = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Psi}^{-1}\boldsymbol{Y}'$. □

Corollary 2.3. *If $\boldsymbol{Y} \sim SNC_{p\times n}(\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$, then*

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto \phi_{p\times n}(\boldsymbol{M}; \boldsymbol{\Lambda}\boldsymbol{\tau}, \boldsymbol{I}_n \otimes \boldsymbol{\Lambda}) F_C\big((\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{I}_n\big),$$

where $\boldsymbol{\tau} = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}$, $\boldsymbol{\Lambda} = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Delta}^{-1})^{-1}$, and $F_C(\cdot; \boldsymbol{I}_n)$ is the CDF of the n-variate standard Cauchy distribution.

Proof. Since the distribution $SNC_{p\times n}(\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$ is in fact the distribution $SSMESN_{p\times n}(\boldsymbol{M}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}, \boldsymbol{I}_n, \boldsymbol{\lambda}, \boldsymbol{0}; (k, s), Q)$ with $k(\theta) = 1$ and $s(\theta, \omega) = \omega^{-\frac{1}{2}}$ which $\omega \sim IGamma(\frac{1}{2}, \frac{1}{2})$, by Proposition 2.1, the posterior density of $\boldsymbol{M}$ is

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto \phi_{p\times n}(\boldsymbol{M}; \boldsymbol{\Lambda}\boldsymbol{\tau}, \boldsymbol{I}_n \otimes \boldsymbol{\Lambda}) \int_{S_G} \Phi_n\big(\omega^{-\frac{1}{2}}(\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{I}_n\big)\, dG(\omega),$$

where G is the CDF of the distribution $IGamma(\frac{1}{2}, \frac{1}{2})$ and S_G is its support. Let $f_C(\cdot; \boldsymbol{I}_n)$ be the PDF of the n-variate standard Cauchy distribution. Since the multivariate Cauchy distribution is a special case of the scale mixture of multivariate normal distributions, we have

$$f_C(\boldsymbol{z}; \boldsymbol{I}_n) = \int_{S_G} \phi_n(\boldsymbol{z}; \boldsymbol{0}, \omega \boldsymbol{I}_n)\, dG(\omega),$$

where $\phi_n(\cdot; \boldsymbol{0}, \omega\boldsymbol{I}_n)$ is the PDF of the n-variate normal distribution $N_n(\boldsymbol{0}, \omega\boldsymbol{I}_n)$. Hence, considering $I(\cdot)$ as the indicator function, we have

$$\begin{aligned}
&\int_{S_G} \Phi_n\big(\omega^{-\frac{1}{2}}(\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{I}_n\big)\, dG(\omega)\\
&= \int_{S_G} \Phi_n\big((\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \omega\boldsymbol{I}_n\big)\, dG(\omega)\\
&= \int_{S_G} \left[\int_{\mathbb{R}^n} I\big(\boldsymbol{z} \leq (\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\big)\phi_n(\boldsymbol{z}; \boldsymbol{0}, \omega\boldsymbol{I}_n) d\boldsymbol{z}\right] dG(\omega)\\
&= \int_{\mathbb{R}^n} I\big(\boldsymbol{z} \leq (\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\big) \left[\int_{S_G} \phi_n(\boldsymbol{z}; \boldsymbol{0}, \omega\boldsymbol{I}_n) dG(\omega)\right] d\boldsymbol{z}\\
&= \int_{\mathbb{R}^n} I\big(\boldsymbol{z} \leq (\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}\big) f_C(\boldsymbol{z}; \boldsymbol{I}_n) d\boldsymbol{z}\\
&= F_C\big((\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{I}_n\big).
\end{aligned}$$

Therefore

$$\pi(\boldsymbol{M}|\boldsymbol{Y}) \propto \phi_{p\times n}(\boldsymbol{M}; \boldsymbol{\Lambda}\boldsymbol{\tau}, \boldsymbol{I}_n \otimes \boldsymbol{\Lambda}) F_C\big((\boldsymbol{Y} - \boldsymbol{M})'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\lambda}; \boldsymbol{I}_n\big).$$

□

3. Some applications

The results obtained in the previous section can be used in many models. In this section, we explain applications of the obtained results in the multivariate linear regression and stress-strength models.

3.1. Multivariate linear regression models

One of the methods of parameter estimation in multivariate linear regression is the Bayesian estimation method. There are many related researches, for example, Arashi, Iranmanesh, Norouzirad, and Salarzadeh Jenatabadi (2014) derived different posterior distributions for the parameters of multivariate regression models with conjugate priors. In this regard, the following corollary presents a different posterior density for the parameters in multivariate linear regression models.

Corollary 3.1. *Consider the p-dimensional vectors* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ *such that*

$$\boldsymbol{x}_i \overset{iid}{\sim} SMN_p(\boldsymbol{B}\boldsymbol{z}_i, \boldsymbol{\Sigma}; k, H), \quad i = 1, \ldots, n, \tag{2}$$

where $\boldsymbol{z}_i$ *is a q-dimensional known vector and* $\boldsymbol{B}$ *is the* $p \times q$ *unknown matrix of regression parameters. If* $\boldsymbol{B}$ *has the prior distribution as* $N_{p\times q}(\mathbf{0}_{p\times q}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \boldsymbol{\Xi})$, *where* $\boldsymbol{Z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)$ *is a* $q \times n$ *known matrix and the* $p \times p$ *known matrix* $\boldsymbol{\Xi}$ *is positive definite, then the posterior density of* $\boldsymbol{B}$ *is given by*

$$\begin{aligned}\pi(\boldsymbol{B}|\underline{\boldsymbol{X}}, \boldsymbol{Z}) \propto \int_{S_H} & etr\{\frac{1}{2k(\theta)}(\boldsymbol{\Pi}_\theta \boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)\underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}\boldsymbol{Z}\underline{\boldsymbol{X}}'\boldsymbol{\Sigma}^{-1}\}|\boldsymbol{\Pi}_\theta|^{\frac{q}{2}} \\ & \times \phi_{p\times q}(\boldsymbol{B}; \boldsymbol{\Pi}_\theta \boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes k(\theta)\boldsymbol{\Pi}_\theta) dH(\theta),\end{aligned}$$

where $\underline{\boldsymbol{X}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)_{p\times n}$ *and* $\boldsymbol{\Pi}_\theta = (\boldsymbol{\Sigma}^{-1} + k(\theta)\boldsymbol{\Xi}^{-1})^{-1}$.

Proof. From (2) it follows that $\underline{\boldsymbol{X}} \sim SMN_{p\times n}(\boldsymbol{B}\boldsymbol{Z}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}; k, H)$ and consequently $\underline{\boldsymbol{X}} \mid \theta = \theta_0 \sim N_{p\times n}(\boldsymbol{B}\boldsymbol{Z}, \boldsymbol{I}_n \otimes k(\theta_0)\boldsymbol{\Sigma})$. By properties of the matrix variate normal distribution, we know

$$\underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1} \sim SMN_{p\times q}(\boldsymbol{B}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \boldsymbol{\Sigma}; k, H).$$

Because of $\boldsymbol{B} \sim N_{p\times q}(\mathbf{0}_{p\times q}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \boldsymbol{\Xi})$, the proof is completed by using Proposition 2.1 with $\boldsymbol{Y}_{p\times q} = \underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}$, $\boldsymbol{\Psi}_{q\times q} = (\boldsymbol{Z}\boldsymbol{Z}')^{-1}$, $\boldsymbol{\Delta} = \boldsymbol{\Xi}$ and $\boldsymbol{\lambda} = \mathbf{0}$. □

By using the posterior density of Corollary 3.1, it is obvious that the Bayes estimator of $\boldsymbol{B}$ under the squared error loss function is

$$\hat{\boldsymbol{B}}_{Bayes} = \int_{\mathbb{R}^{p\times q}} \boldsymbol{B}\pi(\boldsymbol{B}|\underline{\boldsymbol{X}}, \boldsymbol{Z}) d\boldsymbol{B}.$$

The following examples are some special cases of Corollary 3.1.

Example 3.1. If $\boldsymbol{x}_i \overset{iid}{\sim} N_p(\boldsymbol{B}\boldsymbol{z}_i, \boldsymbol{\Sigma})$ for $i = 1, \ldots, n$, then

$$\boldsymbol{B}|\underline{\boldsymbol{X}}, \boldsymbol{Z} \sim N_{p\times q}(\boldsymbol{\Pi}\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \boldsymbol{\Pi}),$$

with $\boldsymbol{\Pi} = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Xi}^{-1})^{-1}$. Therefore, the Bayes estimator of the regression parameters is

$$\hat{\boldsymbol{B}}_{Bayes} = \boldsymbol{\Pi}\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}.$$

Note that because the least squares estimator of $\boldsymbol{B}$ is $\hat{\boldsymbol{B}}_{LS} = \underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}$, we have

$$\hat{\boldsymbol{B}}_{Bayes} = \boldsymbol{\Pi}\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{B}}_{LS}.$$

Example 3.2. If $\boldsymbol{x}_i \overset{iid}{\sim} T_p(\boldsymbol{B}\boldsymbol{z}_i, \boldsymbol{\Sigma}, \nu)$ for $i = 1, \ldots, n$, then $\hat{\boldsymbol{B}}_{Bayes} = \int_{\mathbb{R}^{p\times q}} \boldsymbol{B}\pi(\boldsymbol{B}|\underline{\boldsymbol{X}}, \boldsymbol{Z}) d\boldsymbol{B}$ with

$$\pi(\boldsymbol{B}|\underline{\boldsymbol{X}}, \boldsymbol{Z}) \propto E_\theta\left[etr\{\frac{1}{2\theta}(\boldsymbol{\Pi}_\theta\boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)\boldsymbol{V}\boldsymbol{Z}\underline{\boldsymbol{X}}'\boldsymbol{\Sigma}^{-1}\}|\boldsymbol{\Pi}_\theta|^{\frac{q}{2}}\phi_{p\times q}(\boldsymbol{B}; \boldsymbol{\Pi}_\theta\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \theta\boldsymbol{\Pi}_\theta)\right],$$

where $\theta \sim IGamma(\frac{\nu}{2}, \frac{\nu}{2})$, $\boldsymbol{V} = \underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}$ and $\boldsymbol{\Pi}_\theta = (\boldsymbol{\Sigma}^{-1} + \theta\boldsymbol{\Xi}^{-1})^{-1}$.

3.2. Stress-strength models

A common aspect of research on stress-strength models is the estimation of the reliability of these models. According to this, the Bayesian estimation of stress-strength reliability of elliptically contoured distributions and multivariate skew-normal distribution have been discussed by Kotz, Lumelskii, and Pensky (2003) and Rezaei and Yousefzadeh (2022), respectively.

Obtaining Bayes estimators for the stress-strength reliability is another application of the result of Proposition 2.1. In some stress-strength models, the reliability is considered as the probability $R = P(\boldsymbol{a}'\boldsymbol{x} + \boldsymbol{b}'\boldsymbol{y} + c > 0)$ where $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{y} \in \mathbb{R}^q$ are two independent random vectors and $\boldsymbol{a} \in \mathbb{R}^p$, $\boldsymbol{b} \in \mathbb{R}^q$, and $c \in \mathbb{R}$ are known. Consider the stress-strength model corresponding to the random vectors

$$\boldsymbol{x} \mid (\theta, \omega) \sim ESN_p(\boldsymbol{\mu}_1, k(\theta)\boldsymbol{\Sigma}_1, s(\theta, \omega)\boldsymbol{\lambda}_1, \delta_1), \tag{3}$$

and

$$\boldsymbol{y} \mid (\theta, \omega) \sim ESN_q(\boldsymbol{\mu}_2, k(\theta)\boldsymbol{\Sigma}_2, s(\theta, \omega)\boldsymbol{\lambda}_2, \delta_2), \tag{4}$$

and denote its reliability by R_{SSMESN}.

Suppose that two independent random samples $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ are from the distributions of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. It is obvious that

$$\underline{\boldsymbol{X}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \sim SSMESN_{p\times n}(\boldsymbol{M}_1, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_1, \boldsymbol{I}_n, \boldsymbol{\lambda}_1, \boldsymbol{\delta}_1; (k, s), Q),$$

and

$$\underline{\boldsymbol{Y}} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m) \sim SSMESN_{q\times m}(\boldsymbol{M}_2, \boldsymbol{I}_m \otimes \boldsymbol{\Sigma}_2, \boldsymbol{I}_m, \boldsymbol{\lambda}_2, \boldsymbol{\delta}_2; (k, s), Q),$$

where $\boldsymbol{M}_1 = \boldsymbol{1}'_n \otimes \boldsymbol{\mu}_1$, $\boldsymbol{\delta}_1 = \delta_1 \boldsymbol{1}_n$, $\boldsymbol{M}_2 = \boldsymbol{1}'_m \otimes \boldsymbol{\mu}_2$ and $\boldsymbol{\delta}_2 = \delta_2 \boldsymbol{1}_m$. Let $\boldsymbol{\Sigma}_i$, $\boldsymbol{\lambda}_i$ and $\boldsymbol{\delta}_i$ for $i = 1, 2$ are known and consider the prior distributions $N_{p\times n}(\boldsymbol{0}_{p\times n}, \boldsymbol{I}_n \otimes \boldsymbol{\Delta}_1)$ and $N_{q\times m}(\boldsymbol{0}_{q\times m}, \boldsymbol{I}_m \otimes \boldsymbol{\Delta}_2)$ for $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$, respectively. By Proposition 2.1, the posterior densities of $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$, i.e. $\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}})$ and $\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}})$, have the form (1). Since $\boldsymbol{\mu}_1 = \frac{1}{n}\boldsymbol{M}_1\boldsymbol{1}_n$ and $\boldsymbol{\mu}_2 = \frac{1}{m}\boldsymbol{M}_2\boldsymbol{1}_m$, the reliability R_{SSMESN} is a function of $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$, i.e. $R_{SSMESN}(\boldsymbol{M}_1, \boldsymbol{M}_2)$. Therefore, the Bayes estimator of R_{SSMESN} under the squared error loss function is obtained by

$$\hat{R}_{SSMESN} = \int_{\mathbb{R}^{p\times n}} \int_{\mathbb{R}^{q\times m}} R_{SSMESN}(\boldsymbol{M}_1, \boldsymbol{M}_2)\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}})\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}})d\boldsymbol{M}_2 d\boldsymbol{M}_1. \tag{5}$$

The following examples present the Bayes estimator of the stress-strength reliability of some multivariate distributions such as normal, skew t, and skew-normal-Cauchy.

Example 3.3. By considering $k(\theta) = s(\theta, \omega) = 1$, $\boldsymbol{\lambda}_1 = \boldsymbol{0}$, $\boldsymbol{\lambda}_2 = \boldsymbol{0}$, and $\delta_1 = \delta_2 = 0$ in (3) and (4), the Bayes estimator of the stress-strength reliability corresponding to the multivariate normal distributions, R_N, is obtained by (5) where

$$R_{SSMESN}(\boldsymbol{M}_1, \boldsymbol{M}_2) \equiv R_N(\boldsymbol{M}_1, \boldsymbol{M}_2) = \Phi_1\left(\frac{\frac{1}{n}\boldsymbol{a}'\boldsymbol{M}_1\boldsymbol{1}_n + \frac{1}{m}\boldsymbol{b}'\boldsymbol{M}_2\boldsymbol{1}_m + c}{\sqrt{\boldsymbol{a}'\boldsymbol{\Sigma}_1\boldsymbol{a} + \boldsymbol{b}'\boldsymbol{\Sigma}_2\boldsymbol{b}}}\right),$$

and by Corollary 2.1,

$$\boldsymbol{M}_1|\underline{\boldsymbol{X}} \sim N_{p\times n}(\boldsymbol{\Lambda}_1\boldsymbol{\Sigma}_1^{-1}\underline{\boldsymbol{X}}, \boldsymbol{I}_n \otimes \boldsymbol{\Lambda}_1), \quad \text{and} \quad \boldsymbol{M}_2|\underline{\boldsymbol{Y}} \sim N_{q\times m}(\boldsymbol{\Lambda}_2\boldsymbol{\Sigma}_2^{-1}\underline{\boldsymbol{Y}}, \boldsymbol{I}_m \otimes \boldsymbol{\Lambda}_2)$$

with $\boldsymbol{\Lambda}_i = (\boldsymbol{\Sigma}_i^{-1} + \boldsymbol{\Delta}_i^{-1})^{-1}$ for $i = 1, 2$.

Example 3.4. The reliability R_{SSMESN} in (5) becomes the stress-strength reliability of the multivariate skew t distributions (R_{ST}), which has been calculated by Rezaei and Yousefzadeh (2022), if it is considered

$$\delta_1 = 0,\ k(\theta_0) = \theta_0,\ s(\theta_0, \omega_0) = 1,\ \theta_1 \sim IGamma(\tfrac{\nu_1}{2}, \tfrac{\nu_1}{2}),$$

and

$$\delta_2 = 0,\ k(\theta_0) = \theta_0,\ s(\theta_0, \omega_0) = 1,\ \theta_2 \sim IGamma(\tfrac{\nu_2}{2}, \tfrac{\nu_2}{2}),$$

in (3) and (4), respectively. Hence, the Bayes estimator of R_{ST} is

$$\hat{R}_{ST}^{Bayes} = \int_{\mathbb{R}^{p\times n}} \int_{\mathbb{R}^{q\times m}} R_{ST}(\boldsymbol{M}_1, \boldsymbol{M}_2)\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}})\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}})d\boldsymbol{M}_2 d\boldsymbol{M}_1,$$

where from Corollary 2.2,

$$\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}}) \propto E_{\theta_1}\left[\varrho^1_{\theta_1}|\boldsymbol{\Lambda}^1_{\theta_1}|^{\frac{n}{2}}\phi_{p\times n}(\boldsymbol{M}_1; \boldsymbol{\Lambda}^1_{\theta_1}\boldsymbol{\tau}_1, \boldsymbol{I}_n \otimes \theta_1\boldsymbol{\Lambda}^1_{\theta_1})\Phi_n(\theta_1^{-\frac{1}{2}}(\underline{\boldsymbol{X}} - \boldsymbol{M}_1)'\boldsymbol{\Sigma}_1^{-\frac{1}{2}}\boldsymbol{\lambda}_1; \boldsymbol{I}_n)\right],$$

and

$$\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}}) \propto E_{\theta_2}\left[\varrho^2_{\theta_2}|\boldsymbol{\Lambda}^2_{\theta_2}|^{\frac{m}{2}}\phi_{q\times m}(\boldsymbol{M}_2; \boldsymbol{\Lambda}^2_{\theta_2}\boldsymbol{\tau}_2, \boldsymbol{I}_m \otimes \theta_2\boldsymbol{\Lambda}^2_{\theta_2})\Phi_{n_2}(\theta_2^{-\frac{1}{2}}(\underline{\boldsymbol{Y}} - \boldsymbol{M}_2)'\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\boldsymbol{\lambda}_2; \boldsymbol{I}_m)\right],$$

with $\boldsymbol{\Lambda}^i_{\theta_i} = (\boldsymbol{\Sigma}_i^{-1} + \theta_i\boldsymbol{\Delta}_i^{-1})^{-1}$ for $i = 1, 2$, $\boldsymbol{\tau}_1 = \boldsymbol{\Sigma}_1^{-1}\underline{\boldsymbol{X}}$, $\boldsymbol{\tau}_2 = \boldsymbol{\Sigma}_2^{-1}\underline{\boldsymbol{Y}}$, $\varrho^1_{\theta_1} = etr\left\{\frac{\boldsymbol{\Lambda}^1_{\theta_1}\boldsymbol{\tau}_1\boldsymbol{\tau}_1' - \boldsymbol{\tau}_1\underline{\boldsymbol{X}}'}{2\theta_1}\right\}$ and $\varrho^2_{\theta_2} = etr\left\{\frac{\boldsymbol{\Lambda}^2_{\theta_2}\boldsymbol{\tau}_2\boldsymbol{\tau}_2' - \boldsymbol{\tau}_2\underline{\boldsymbol{Y}}'}{2\theta_2}\right\}$.

Example 3.5. The distribution of the random vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are $SNC_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\lambda}_1)$ and $SNC_q(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \boldsymbol{\lambda}_2)$, respectively, by considering $\delta_1 = \delta_2 = 0$, $k(\theta_0) = \theta_0$ and $s(\theta_0, \omega_0) = \omega_0^{-\frac{1}{2}}$ in (3) and (4) when $\omega \sim IGamma(\frac{1}{2}, \frac{1}{2})$. From (5), the Bayes estimator of the stress-strength reliability of these vectors (R_{SNC}) is given by

$$\hat{R}_{SNC}^{Bayes} = \int_{\mathbb{R}^{p\times n}} \int_{\mathbb{R}^{q\times m}} R_{SNC}(\boldsymbol{M}_1, \boldsymbol{M}_2)\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}})\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}})d\boldsymbol{M}_2 d\boldsymbol{M}_1, \tag{6}$$

where by Corollary 2.3,

$$\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}}) \propto \phi_{p\times n}(\boldsymbol{M}_1; \boldsymbol{\Lambda}_1\boldsymbol{\tau}_1, \boldsymbol{I}_n \otimes \boldsymbol{\Lambda}_1)F_C((\underline{\boldsymbol{X}} - \boldsymbol{M}_1)'\boldsymbol{\Sigma}_1^{-\frac{1}{2}}\boldsymbol{\lambda}_1; \boldsymbol{I}_n),$$

$$\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}}) \propto \phi_{q\times m}(\boldsymbol{M}_2; \boldsymbol{\Lambda}_2\boldsymbol{\tau}_2, \boldsymbol{I}_m \otimes \boldsymbol{\Lambda}_2)F_C((\underline{\boldsymbol{Y}} - \boldsymbol{M}_2)'\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\boldsymbol{\lambda}_2; \boldsymbol{I}_m),$$

with $\boldsymbol{\tau}_1 = \boldsymbol{\Sigma}_1^{-1}\underline{\boldsymbol{X}}$, $\boldsymbol{\tau}_2 = \boldsymbol{\Sigma}_2^{-1}\underline{\boldsymbol{Y}}$ and $\boldsymbol{\Lambda}_i = (\boldsymbol{\Sigma}_i^{-1} + \boldsymbol{\Delta}_i^{-1})$ for $i = 1, 2$, and also by Rezaei and Yousefzadeh (2022),

$$\begin{aligned} R_{SNC}(\boldsymbol{M}_1, \boldsymbol{M}_2) =& R_N(\boldsymbol{M}_1, \boldsymbol{M}_2) + \frac{1}{\pi}\Bigg[\int_0^\infty \frac{\cos\left((\frac{1}{n}\boldsymbol{a}'\boldsymbol{M}_1\boldsymbol{1}_n + \frac{1}{m}\boldsymbol{b}'\boldsymbol{M}_2\boldsymbol{1}_m + c)u\right)}{u} \\ & \times e^{\frac{-u^2}{2}(\boldsymbol{a}'\boldsymbol{\Sigma}_1\boldsymbol{a} + \boldsymbol{b}'\boldsymbol{\Sigma}_2\boldsymbol{b})}\left(\overset{*}{\tau}_{\boldsymbol{\Sigma}_1,\boldsymbol{\lambda}_1}(\boldsymbol{a}u) + \overset{*}{\tau}_{\boldsymbol{\Sigma}_2,\boldsymbol{\lambda}_2}(\boldsymbol{b}u)\right)du \\ & - \int_0^\infty \frac{\sin\left((\frac{1}{n}\boldsymbol{a}'\boldsymbol{M}_1\boldsymbol{1}_n + \frac{1}{m}\boldsymbol{b}'\boldsymbol{M}_2\boldsymbol{1}_m + c)u\right)}{u} \\ & \times e^{\frac{-u^2}{2}(\boldsymbol{a}'\boldsymbol{\Sigma}_1\boldsymbol{a} + \boldsymbol{b}'\boldsymbol{\Sigma}_2\boldsymbol{b})}\,\overset{*}{\tau}_{\boldsymbol{\Sigma}_1,\boldsymbol{\lambda}_1}(\boldsymbol{a}u)\,\overset{*}{\tau}_{\boldsymbol{\Sigma}_2,\boldsymbol{\lambda}_2}(\boldsymbol{b}u)du\Bigg], \end{aligned}$$

where $\overset{*}{\tau}_{\boldsymbol{\Sigma},\boldsymbol{\lambda}}(\boldsymbol{t}) = \int_0^\infty \tau\left(\frac{\boldsymbol{\lambda}'\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{t}}{\sqrt{1+x^2\boldsymbol{\lambda}'\boldsymbol{\lambda}}}x\right)\phi_1(x)dx$ with the PDF of the univariate standard normal distribution ϕ_1 and $\tau(z) = \sqrt{\frac{2}{\pi}}\int_0^z \exp\{\frac{t^2}{2}\}dt$.

4. A simulation study

In this section, a simulation study is presented to compare different Bayes estimators of R_{SNC}, the stress-strength reliability of the multivariate skew normal-Cauchy distributions. Note here that all relevant programs are written in the R software package and are performed by using a machine equipped with an Intel Core i5-3230M 2.60 GHz processor and 4 GB RAM. The R codes can be obtained on request from the authors.

In this simulation study, we focus on comparing the Bayes estimations of the stress-strength reliability corresponding to the random vectors

$$\boldsymbol{x} \sim SNC_3\left(\boldsymbol{\mu}_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \boldsymbol{\Sigma}_1 = \begin{pmatrix} 4.5 & 1.5 & -0.4 \\ 1.5 & 3.0 & 2.3 \\ -0.4 & 2.3 & 4.5 \end{pmatrix}, \boldsymbol{\lambda}_1 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}\right),$$

and

$$\boldsymbol{y} \sim SNC_3\left(\boldsymbol{\mu}_2 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \boldsymbol{\Sigma}_2 = \begin{pmatrix} 2.0 & -0.5 & -0.8 \\ -0.5 & 2.2 & -0.3 \\ -0.8 & -0.3 & 1.8 \end{pmatrix}, \boldsymbol{\lambda}_2 = \begin{pmatrix} 0.3 \\ -0.3 \\ -0.3 \end{pmatrix}\right),$$

with $\boldsymbol{a} = \boldsymbol{b} = (0.25, 0.5, -0.25)'$ and $c = 1$ which equals to 0.82164.

We have taken the prior distributions $N_{3\times n}(\mathbf{0}_{3\times n}, \boldsymbol{I}_n \otimes \boldsymbol{\Delta}_1)$ and $N_{3\times m}(\mathbf{0}_{3\times m}, \boldsymbol{I}_m \otimes \boldsymbol{\Delta}_2)$ to obtain the Bayes estimations of R_{SNC}, where

- **Prior-1**: $\boldsymbol{\Delta}_1 = \boldsymbol{I}_3$ and $\boldsymbol{\Delta}_2 = \boldsymbol{I}_3$;
- **Prior-2**: $\boldsymbol{\Delta}_1 = \begin{pmatrix} 2.73 & -0.66 & -1.59 \\ -0.66 & 2.73 & 1.35 \\ -1.59 & 1.35 & 2.73 \end{pmatrix}$, and $\boldsymbol{\Delta}_2 = \begin{pmatrix} 1.66 & 1.66 & 1.42 \\ 1.66 & 3.55 & 1.66 \\ 1.42 & 1.66 & 2.81 \end{pmatrix}$;
- **Prior-3**: $\boldsymbol{\Delta}_1 = \begin{pmatrix} 1.66 & 1.66 & 1.42 \\ 1.66 & 3.55 & 1.66 \\ 1.42 & 1.66 & 2.81 \end{pmatrix}$, and $\boldsymbol{\Delta}_2 = \begin{pmatrix} 2.73 & -0.66 & -1.59 \\ -0.66 & 2.73 & 1.35 \\ -1.59 & 1.35 & 2.73 \end{pmatrix}$.

As can be seen in (6) the Bayes estimators of R_{SNC} have not closed forms; hence we use Markov Chain Monte Carlo integration to calculate them. For generating random samples from the posterior densities $\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}})$ and $\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}})$, we employ the independence sampler, a special case of the Metropolis-Hastings sampler, as follows; see Tierney (1994) for more.

Algorithm 1. Perform the following steps to generate random sample from the posterior density $\pi(\boldsymbol{M}|\boldsymbol{Y})$ of Corollary 2.3:

- Step 1: Generate the initial matrix $\boldsymbol{M}_0$ from $N_{p\times n}(\mathbf{0}_{p\times n}, \boldsymbol{I}_n \otimes \boldsymbol{\Delta})$.
- Step 2: Calculate the matrices $\boldsymbol{\tau} = \boldsymbol{\Sigma}^{-1}\boldsymbol{Y}$, $\boldsymbol{\Sigma}_g = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Delta}^{-1})^{-1}$ and $\boldsymbol{M}_g = \boldsymbol{\Sigma}_g\boldsymbol{\tau}$.
- Step 3: Generate $\boldsymbol{M}_\pi$ and U from $N_{p\times n}(\boldsymbol{M}_g, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_g)$ and $Uniform(0, 1)$, respectively.
- Step 4: If
$$U < \frac{\pi(\boldsymbol{M}_\pi|\boldsymbol{Y})\phi_{p\times n}(\boldsymbol{M}_0; \boldsymbol{M}_g, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_g)}{\pi(\boldsymbol{M}_0|\boldsymbol{Y})\phi_{p\times n}(\boldsymbol{M}_\pi; \boldsymbol{M}_g, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_g)},$$
accept $\boldsymbol{M}_\pi$ and consider it as a random sample of $\pi(\boldsymbol{M}|\boldsymbol{Y})$; otherwise, set $\boldsymbol{M}_\pi$ as $\boldsymbol{M}_0$ and repeat steps 3 and 4.

For monitoring convergence of the algorithm 1, we use the Gelman-Rubin diagnostic which is a method to check the convergence of a Metropolis-Hastings chain, see Gelman and Rubin (1992). For this method, we run 10 independent chains with different initial values and discard the first 20% of generated samples as the burn-in step in each chain. Table 1 contains the Gelman-Rubin statistic for different posterior densities $\pi(\boldsymbol{M}_1|\underline{\boldsymbol{X}})$ and $\pi(\boldsymbol{M}_2|\underline{\boldsymbol{Y}})$. These values are close to 1 that indicate the convergence of Algorithm 1 for each posterior density.

Bias and mean squared error (MSE) of different Bayes estimations of R_{SNC} for varying sample sizes are presented in Table 2. It can be observed that for any pair (n, m), the absolute value

Table 1: The Gelman-Rubin statistic of different posterior densities

	Prior-1	Prior-2	Prior-3
$\pi(\boldsymbol{M}_1\vert\underline{\boldsymbol{X}})$	0.9897789	0.9897852	0.9898353
$\pi(\boldsymbol{M}_2\vert\underline{\boldsymbol{Y}})$	0.9897716	0.9898104	0.9897663

Table 2: The results of simulation for different Bayes estimations of R_{SNC}

(n, m)	Prior-1		Prior-2		Prior-3	
	Bias	MSE	Bias	MSE	Bias	MSE
$(10, 10)$	-0.076379	0.007716	-0.063564	0.010236	0.022696	0.001999
$(10, 15)$	-0.041828	0.004289	-0.075475	0.010454	0.030584	0.002575
$(10, 20)$	-0.053089	0.004330	-0.065466	0.007348	0.017873	0.001680
$(15, 10)$	-0.044426	0.004964	-0.073285	0.008810	0.011459	0.002159
$(15, 15)$	-0.049522	0.004271	-0.057152	0.007266	0.034053	0.003166
$(15, 20)$	-0.048657	0.003523	-0.067342	0.007193	0.005173	0.002467
$(20, 10)$	-0.039217	0.004371	-0.044141	0.007374	0.025004	0.001743
$(20, 15)$	-0.032857	0.002021	-0.066391	0.007225	0.013092	0.001603
$(20, 20)$	-0.042048	0.002895	-0.061840	0.006779	0.001927	0.001908

of bias and MSE of the Bayes estimation obtained by Prior-3 are lower than that of other Bayes estimations. Hence, Prior-3 introduces considerably more prior information for R_{SNC}. Also, as the sample sizes increase, the MSE of all Bayes estimations decreases.

5. Real data analysis

In this section, we consider Chemical Reaction Data in Box and Youle (1955) and obtain different Bayes estimations for the regression parameters. These data consist of the results of a planned experiment involving a chemical reaction that are given in Table 3. The dependent variables are the percentage of unchanged starting material (x_1), the percentage converted to the desired product (x_2) and the percentage of unwanted by-product (x_3) and also the independent variables are the temperature (z_1), the concentration (z_2) and the time (z_3).

Let $\underline{\boldsymbol{X}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{19})$ and $\boldsymbol{Z} = \begin{pmatrix} 1 & \ldots & 1 \\ \boldsymbol{z}_1 & \ldots & \boldsymbol{z}_{19} \end{pmatrix}$. Consider the regression of (x_1, x_2, x_3) on (z_1, z_2, z_3) in the multivariate normal distribution case. The least squares estimation of the regression parameters

$$\boldsymbol{B} = (\beta_{ij}) = \begin{pmatrix} \beta_{10} & \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{20} & \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{30} & \beta_{31} & \beta_{32} & \beta_{33} \end{pmatrix}.$$

is

$$\hat{\boldsymbol{B}}_{LS} = \underline{\boldsymbol{X}}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1} = \begin{pmatrix} 332.11098 & -1.54596 & -1.42455 & -2.23736 \\ -26.03526 & 0.40455 & 0.29299 & 1.03380 \\ -164.07894 & 0.91392 & 0.89947 & 1.15348 \end{pmatrix}.$$

To test overall regression, we must consider the hypothesis that none of (z_1, z_2, z_3) predict any of (x_1, x_2, x_3). For this purpose, we consider Wilks' Λ and calculate it as

$$\Lambda = \frac{|\underline{\boldsymbol{X}}\,\underline{\boldsymbol{X}}' - \hat{\boldsymbol{B}}_{LS}\boldsymbol{Z}\underline{\boldsymbol{X}}'|}{|\underline{\boldsymbol{X}}\,\underline{\boldsymbol{X}}' - 19\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'|} = 0.03315764.$$

Here, the corresponding lower critical value is $\Lambda_{0.05,3,3,15} = 0.309$. Since $\Lambda < \Lambda_{0.05,3,3,15}$, we reject the hypothesis and hence at least one of (z_1, z_2, z_3) predicts x's. For more information

Table 3: Chemical reaction data

Experiment Number	Dependent variables $\boldsymbol{x}' = (x_1, x_2, x_3)$	Independent variables $\boldsymbol{z}' = (z_1, z_2, z_3)$
1	$\boldsymbol{x}_1' = (41.5, 45.9, 11.2)$	$\boldsymbol{z}_1' = (162, 23, 3)$
2	$\boldsymbol{x}_2' = (33.8, 53.3, 11.2)$	$\boldsymbol{z}_2' = (162, 23, 8)$
3	$\boldsymbol{x}_3' = (27.7, 57.5, 12.7)$	$\boldsymbol{z}_3' = (162, 30, 5)$
4	$\boldsymbol{x}_4' = (21.7, 58.8, 16.0)$	$\boldsymbol{z}_4' = (162, 30, 8)$
5	$\boldsymbol{x}_5' = (19.9, 60.6, 16.2)$	$\boldsymbol{z}_5' = (172, 25, 5)$
6	$\boldsymbol{x}_6' = (15.0, 58.0, 22.6)$	$\boldsymbol{z}_6' = (172, 25, 8)$
7	$\boldsymbol{x}_7' = (12.2, 58.6, 24.5)$	$\boldsymbol{z}_7' = (172, 30, 5)$
8	$\boldsymbol{x}_8' = (4.3, 52.4, 38.0)$	$\boldsymbol{z}_8' = (172, 30, 8)$
9	$\boldsymbol{x}_9' = (19.3, 56.9, 21.3)$	$\boldsymbol{z}_9' = (167, 27.5, 6.5)$
10	$\boldsymbol{x}_{10}' = (6.4, 55.4, 30.8)$	$\boldsymbol{z}_{10}' = (177, 27.5, 6.5)$
11	$\boldsymbol{x}_{11}' = (37.6, 46.9, 14.7)$	$\boldsymbol{z}_{11}' = (157, 27.5, 6.5)$
12	$\boldsymbol{x}_{12}' = (18.0, 57.3, 22.2)$	$\boldsymbol{z}_{12}' = (167, 32.5, 6.5)$
13	$\boldsymbol{x}_{13}' = (26.3, 55.0, 18.3)$	$\boldsymbol{z}_{13}' = (167, 22.5, 6.5)$
14	$\boldsymbol{x}_{14}' = (9.9, 58.9, 28.0)$	$\boldsymbol{z}_{14}' = (167, 27.5, 9.5)$
15	$\boldsymbol{x}_{15}' = (25.0, 50.3, 22.1)$	$\boldsymbol{z}_{15}' = (167, 27.5, 3.5)$
16	$\boldsymbol{x}_{16}' = (14.1, 61.1, 23.0)$	$\boldsymbol{z}_{16}' = (177, 20, 6.5)$
17	$\boldsymbol{x}_{17}' = (15.2, 62.9, 20.7)$	$\boldsymbol{z}_{17}' = (177, 20, 6.5)$
18	$\boldsymbol{x}_{18}' = (15.9, 60.0, 22.1)$	$\boldsymbol{z}_{18}' = (160, 34, 7.5)$
19	$\boldsymbol{x}_{19}' = (19.6, 60.6, 19.3)$	$\boldsymbol{z}_{19}' = (160, 34, 7.5)$

about Wilks' Λ, see Rencher (2002). To determine whether each of the predictors z_1, z_2, and z_3 contributes to the model, we must consider the hypothesis that (x_1, x_2, x_3) do not depend on z_r for $r = 1, 2, 3$. For this purpose, we use Pillai's test statistic which is calculated as

$$V^{(r)} = tr\left((\underline{\boldsymbol{X}}\,\underline{\boldsymbol{X}}' - \hat{\boldsymbol{B}}_{LS}^{r} \boldsymbol{Z}_r \underline{\boldsymbol{X}}')^{-1} (\hat{\boldsymbol{B}}_{LS} \boldsymbol{Z} \underline{\boldsymbol{X}}' - \hat{\boldsymbol{B}}_{LS}^{r} \boldsymbol{Z}_r \underline{\boldsymbol{X}}') \right),$$

where $\hat{\boldsymbol{B}}_{LS}^{r}$ is the least squares estimation of the regression parameters without considering z_r in the model and $\boldsymbol{Z}_r$ contains the rows of $\boldsymbol{Z}$ corresponding to $\hat{\boldsymbol{B}}_{LS}^{r}$. For more information about Pillai's test statistic, see Rencher (2002). Table 4 contains the values of Pillai's test statistic, their approximate F-statistics, and the related p-values. Based on the obtained results, it can be concluded that z_1, z_2, and z_3 are significant predictors for x's.

Table 4: The obtained results for Pillai's test statistics

Variable	Test statistic	Approx. F-statistic	df	p-value
z_1	0.9536407	89.140	(3,13)	6.364×10^{-9}
z_2	0.8881265	34.401	(3,13)	1.894×10^{-6}
z_3	0.7639332	14.023	(3,13)	0.0002279

Hence, by considering $\hat{\boldsymbol{B}}_{LS}$, the fitted model is

$$\hat{\boldsymbol{x}}^{LS} = \begin{pmatrix} \hat{x}_1^{LS} \\ \hat{x}_2^{LS} \\ \hat{x}_3^{LS} \end{pmatrix} = \begin{pmatrix} 332.11098 - 1.54596(z_1) - 1.42455(z_2) - 2.23736(z_3) \\ -26.03526 + 0.40455(z_1) + 0.29299(z_2) + 1.03380(z_3) \\ -164.07894 + 0.91392(z_1) + 0.89947(z_2) + 1.15348(z_3) \end{pmatrix}.$$

Assume the prior distributions $N_{3\times 4}(\boldsymbol{0}_{3\times 4}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \Xi_1)$ and $N_{3\times 4}(\boldsymbol{0}_{3\times 4}, (\boldsymbol{Z}\boldsymbol{Z}')^{-1} \otimes \Xi_2)$ with

$$\Xi_1 = \begin{pmatrix} 9 & 1.5 & 0.5 \\ 1.5 & 6.25 & 0.25 \\ 0.5 & 0.25 & 9 \end{pmatrix} \quad \text{and} \quad \Xi_2 = \begin{pmatrix} 2.25 & 0.05 & -0.25 \\ 0.05 & 2.25 & -0.25 \\ -0.25 & -0.25 & 2.25 \end{pmatrix},$$

for $\boldsymbol{B}$. From Example 3.1,

$$\hat{\boldsymbol{B}}_{Bayes-1} = \left(\hat{\beta}_{ij}^{Bayes-1}\right) = \boldsymbol{\Pi}_1 \boldsymbol{S}^{-1} \hat{\boldsymbol{B}}_{LS} = \begin{pmatrix} 50.93274 & -0.09489 & -0.09204 & -0.05708 \\ 29.22740 & 0.00049 & -0.01894 & 0.13893 \\ 21.08979 & 0.03507 & 0.04355 & 0.06247 \end{pmatrix},$$

and

$$\hat{\boldsymbol{B}}_{Bayes-2} = \left(\hat{\beta}_{ij}^{Bayes-2}\right) = \boldsymbol{\Pi}_2 \boldsymbol{S}^{-1} \hat{\boldsymbol{B}}_{LS} = \begin{pmatrix} 31.22191 & -0.03787 & -0.03720 & 0.00107 \\ 27.46278 & -0.00687 & -0.01644 & 0.08360 \\ 24.01381 & 0.00142 & 0.00409 & 0.03963 \end{pmatrix},$$

where

$$\boldsymbol{S} = \frac{1}{19-1} \sum_{i=1}^{19} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})' = \begin{pmatrix} 99.29620 & -28.56863 & -59.02813 \\ -28.56863 & 22.24801 & 5.78395 \\ -59.02813 & 5.78395 & 45.27140 \end{pmatrix},$$

is the sample covariance matrix of the dependent variables,

$$\boldsymbol{\Pi}_1 = (\boldsymbol{S}^{-1} + \boldsymbol{\Xi}_1^{-1})^{-1} = \begin{pmatrix} 5.16859 & -1.07754 & -2.95790 \\ -1.07754 & 3.50550 & -1.67304 \\ -2.95790 & -1.67304 & 4.38077 \end{pmatrix},$$

and

$$\boldsymbol{\Pi}_2 = (\boldsymbol{S}^{-1} + \boldsymbol{\Xi}_2^{-1})^{-1} = \begin{pmatrix} 1.74444 & -0.41109 & -0.76121 \\ -0.41109 & 1.62515 & -0.66010 \\ -0.76121 & -0.66010 & 1.58611 \end{pmatrix}.$$

To examine the significance of $\hat{\boldsymbol{B}}_{Bayes-1}$ and $\hat{\boldsymbol{B}}_{Bayes-2}$, we use the following algorithm for obtaining bootstrap confidence intervals for their elements. See Efron and Tibshirani (1994) for more information about the bootstrapping methods.

Algorithm 2. Suppose that $\hat{\boldsymbol{B}}_{Bayes} = (\hat{\beta}_{ij})$ with $i = 1, 2, 3$ and $j = 0, 1, 2, 3$, is a Bayes estimation for $\boldsymbol{B}$ and $\boldsymbol{z}_i^* = (1, \boldsymbol{z}_i)'$ for $i = 1, \dots, 19$.

- Step 1: Generate $\boldsymbol{x}_l^*$ from $N_p(\hat{\boldsymbol{B}}_{Bayes} \boldsymbol{z}_l^*, \boldsymbol{S})$ for $l = 1, \dots, 19$.
- Step 2: Obtain the Bayes estimation $\hat{\boldsymbol{B}}_{Bayes}^{(1)} = \left(\hat{\beta}_{ij}^{(1)}\right)$ based on $\boldsymbol{x}_1^*, \dots, \boldsymbol{x}_{19}^*$.
- Step 3: Repeat the above steps as much as T repetition until get the parametric bootstrap sample $\hat{\boldsymbol{B}}_{Bayes}^{(1)} = \left(\hat{\beta}_{ij}^{(1)}\right), \dots, \hat{\boldsymbol{B}}_{Bayes}^{(T)} = \left(\hat{\beta}_{ij}^{(T)}\right)$.
- Step 4: Obtain a $100(1-\alpha)\%$ percentile bootstrap confidence interval for $\hat{\beta}_{ij}$ by $\left(\hat{\beta}_{ij}^{\frac{\alpha}{2}}, \hat{\beta}_{ij}^{1-\frac{\alpha}{2}}\right)$, where $\hat{\beta}_{ij}^{\frac{\alpha}{2}}$ and $\hat{\beta}_{ij}^{1-\frac{\alpha}{2}}$ are respectively $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ percentiles of the parametric bootstrap sample $\hat{\beta}_{ij}^{(1)}, \dots, \hat{\beta}_{ij}^{(T)}$.

Note that falling each element of $\hat{\boldsymbol{B}}_{Bayes}$ within its obtained confidence interval indicates its significance.

The bootstrap 95% confidence intervals (considering $T = 5000$) for the elements of $\hat{\boldsymbol{B}}_{Bayes-1}$ and $\hat{\boldsymbol{B}}_{Bayes-2}$ are presented in Tables 5 and 6.

The results of Table 5 show that among the elements of $\hat{\boldsymbol{B}}_{Bayes-1}$, $\hat{\beta}_{11}^{Bayes-1}$ and $\hat{\beta}_{10}^{Bayes-1}$ are not significant, while the results of Table 6 show the significance of all elements of $\hat{\boldsymbol{B}}_{Bayes-2}$. Therefore, the fitted models are as follows:

$$\hat{\boldsymbol{x}}^{Bayes-1} = \begin{pmatrix} \hat{x}_1^{Bayes-1} \\ \hat{x}_2^{Bayes-1} \\ \hat{x}_3^{Bayes-1} \end{pmatrix} = \begin{pmatrix} -0.09204(z_2) - 0.05708(z_3) \\ 29.22740 + 0.00049(z_1) - 0.01894(z_2) + 0.13893(z_3) \\ 21.08979 + 0.03507(z_1) + 0.04355(z_2) + 0.06247(z_3) \end{pmatrix}$$

Table 5: The bootstrap confidence intervals of $\hat{\boldsymbol{B}}_{Bayes-1}$'s elements

Element	Estimation	Confidence interval		Explanation
		0.025 percentile	0.975 percentile	
$\hat{\beta}_{10}^{Bayes-1}$	50.93274	24.02470	45.34254	Not significant
$\hat{\beta}_{20}^{Bayes-1}$	29.22740	11.90661	40.36312	Significant
$\hat{\beta}_{20}^{Bayes-1}$	21.08979	18.54900	47.23474	Significant
$\hat{\beta}_{11}^{Bayes-1}$	-0.09489	-0.07655	0.03598	Not significant
$\hat{\beta}_{21}^{Bayes-1}$	0.00049	-0.09020	0.06050	Significant
$\hat{\beta}_{21}^{Bayes-1}$	0.03507	-0.08981	0.06268	Significant
$\hat{\beta}_{12}^{Bayes-1}$	-0.09204	-0.10725	0.05836	Significant
$\hat{\beta}_{22}^{Bayes-1}$	-0.01894	-0.12630	0.09132	Significant
$\hat{\beta}_{22}^{Bayes-1}$	0.04355	-0.12706	0.10032	Significant
$\hat{\beta}_{13}^{Bayes-1}$	-0.05708	-0.14867	0.23455	Significant
$\hat{\beta}_{23}^{Bayes-1}$	0.13893	-0.19981	0.32302	Significant
$\hat{\beta}_{23}^{Bayes-1}$	0.06247	-0.22406	0.30187	Significant

Table 6: The bootstrap confidence intervals of $\hat{\boldsymbol{B}}_{Bayes-2}$'s elements

Element	Estimation	Confidence interval		Explanation
		0.025 percentile	0.975 percentile	
$\hat{\beta}_{10}^{Bayes-2}$	31.22191	19.82957	31.25661	Significant
$\hat{\beta}_{20}^{Bayes-2}$	27.46278	17.69807	33.14585	Significant
$\hat{\beta}_{20}^{Bayes-2}$	24.01381	19.74956	34.28943	Significant
$\hat{\beta}_{11}^{Bayes-2}$	-0.03787	-0.04425	0.01648	Significant
$\hat{\beta}_{21}^{Bayes-2}$	-0.00687	-0.05500	0.02735	Significant
$\hat{\beta}_{21}^{Bayes-2}$	0.00142	-0.05218	0.02614	Significant
$\hat{\beta}_{12}^{Bayes-2}$	-0.03720	-0.06016	0.02834	Significant
$\hat{\beta}_{22}^{Bayes-2}$	-0.01644	-0.07536	0.04321	Significant
$\hat{\beta}_{22}^{Bayes-2}$	0.00409	-0.07124	0.04358	Significant
$\hat{\beta}_{13}^{Bayes-2}$	0.00107	-0.06942	0.13907	Significant
$\hat{\beta}_{23}^{Bayes-2}$	0.08360	-0.09388	0.18981	Significant
$\hat{\beta}_{23}^{Bayes-2}$	0.03963	-0.09804	0.16986	Significant

$$\hat{\boldsymbol{x}}^{Bayes-2} = \begin{pmatrix} \hat{x}_1^{Bayes-2} \\ \hat{x}_2^{Bayes-2} \\ \hat{x}_3^{Bayes-2} \end{pmatrix} = \begin{pmatrix} 31.22191 - 0.03787(z_1) - 0.03720(z_2) + 0.00107(z_3) \\ 27.46278 - 0.00687(z_1) - 0.01644(z_2) + 0.08360(z_3) \\ 24.01381 + 0.00142(z_1) + 0.00409(z_2) + 0.03963(z_3) \end{pmatrix}$$

To check the distribution of the residuals, we use Royston's multivariate normality test; see Royston (1983) for further information. The results of Royston's tests, which are presented in Table 7, show the multivariate normality of the residuals of the fitted models.

Conclusion

In this paper, to obtain a Bayes estimator for the mean matrix of the matrix variate SSMESN distributions, we presented a posterior density by considering a matrix variate normal distribution as its prior. As applications of the results, we used the obtained posterior density for estimating in the multivariate linear regression and the stress-strength models. Finally, different Bayes estimations of the stress-strength reliability of the multivariate skew normal-Cauchy distributions were compared by a simulation study, and the Bayes estimations of

Table 7: The results of Royston's tests

Model	Test statistic	p-value
$\hat{\boldsymbol{x}}^{LS}$	1.192189	0.5718729
$\hat{\boldsymbol{x}}^{Bayes-1}$	2.334849	0.2095501
$\hat{\boldsymbol{x}}^{Bayes-2}$	2.206623	0.2177015

the regression parameters for Chemical Reaction Data were obtained. Here, we focused on obtaining a Bayes estimator for the mean matrix of the matrix variate SSMESN distributions when the matrix variate normal distribution was considered as prior and other parameters of this family of matrix variate distributions were known. In this way, obtaining Bayes estimators for the mean matrix by considering other priors for it, obtaining Bayes estimators for other parameters, using other methods to estimate parameters of this family of distributions, comparing the Bayesian estimators of the parameters with other types of their estimators and applying artificial intelligence algorithms in this regard can be considered as future works.

Acknowledgments

The authors would like to thank the Editor-In-Chief Prof. Dr. Matthias Templ and two anonymous referees for their valuable suggestions and comments on improving the paper.

References

Arashi M, Iranmanesh A, Norouzirad M, Salarzadeh Jenatabadi H (2014). "Bayesian Analysis in Multivariate Regression Models with Conjugate Priors." *Statistics*, **48**(6), 1324–1334. doi:10.1080/02331888.2013.809720.

Box GEP, Youle PV (1955). "The Exploration and Exploitation of Response Surfaces: An Example of the Link between the Fitted Surface and the Basic Mechanism of the System." *Biometrics*, **11**(3), 287–323. doi:10.2307/3001769.

Efron B, Tibshirani RJ (1994). *An Introduction to the Bootstrap*. CRC press.

Gelman A, Rubin DB (1992). "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science*, **7**(4), 457–472. doi:10.1214/ss/1177011136.

Gupta AK, Nagar DK (1999). *Matrix Variate Distributions*, volume 104. CRC Press.

Gupta AK, Varga T (1995). "Normal Mixture Representations of Matrix Variate Elliptically Contoured Distributions." *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 68–78.

Gupta AK, Varga T, Bodnar T (2013). *Elliptically Contoured Models in Statistics and Portfolio Theory*. Springer Science & Business Media.

Harrar SW, Gupta AK (2008). "On Matrix Variate Skew-normal Distributions." *Statistics*, **42**(2), 179–194. doi:10.1080/02331880701597339.

Kotz S, Lumelskii Y, Pensky M (2003). *The Stress–strength Model and Its Generalizations: Theory and Applications*. World Scientific. doi:10.1142/9789812564511_0001.

Ning W, Gupta AK (2012). "Matrix Variate Extended Skew Normal Distributions." *Random Operators and Stochastic Equations*, **20**(4), 299–310. doi:10.1515/rose-2012-0014.

Rencher AC (2002). *Methods of Multivariate Analysis*. A John Wiley & Sons. Inc. Publication. doi:10.1002/0471271357.

Rezaei A, Yousefzadeh F (2022). "Stress-strength Reliability of Generalized Skew-elliptical Distributions and Its Bayes Estimation." *Communications in Statistics-Simulation and Computation*, **51**(7), 3471–3486. doi:10.1080/03610918.2020.1732420.

Rezaei A, Yousefzadeh F, Arellano-Valle RB (2020). "Scale and Shape Mixtures of Matrix Variate Extended Skew Normal Distributions." *Journal of Multivariate Analysis*, **179**, 104649. doi:10.1016/j.jmva.2020.104649.

Royston JP (1983). "Some Techniques for Assessing Multivarate Normality Based on the Shapiro-Wilk W." *Journal of the Royal Statistical Society Series C: Applied Statistics*, **32**(2), 121–133. doi:10.2307/2347291.

Tierney L (1994). "Markov Chains for Exploring Posterior Distributions." *The Annals of Statistics*, pp. 1701–1728.

Affiliation:

Amir Rezaei
Ph.D. graduate in Statistics from the University of Birjand
Birjand, Iran
E-mail: amir.rezaei.st@gmail.com

Fatemeh Yousefzadeh
Department of Statistics
University of Birjand
Birjand, Iran
E-mail: fyousefzadeh@birjand.ac.ir

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 53 *Submitted:* 2022-11-07
2024 *Accepted:* 2023-09-26