Iranian Journal of Mathematical Sciences and Informatics
Vol. 5, No. 2 (2010), pp 33-43
DOI: 10.7508/ijmsi.2010.02.004

On Conditional Applications of Matrix Variate Normal Distribution

Anis Iranmanesh[a,*], M. Arashi[b] and S. M. M. Tabatabaey[a]

[a]Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran
[b]Department of Mathematics, Shahrood University of Technology, P.O. Box 316-3619995161, Shahrood, Iran

E-mail: anisiranmanesh@yahoo.com

Abstract. In this paper, by conditioning on the matrix variate normal distribution (MVND) the construction of the matrix t-type family is considered, thus providing a new perspective of this family. Some important statistical characteristics are given. The presented t-type family is an extension to the work of Dickey [8]. A Bayes estimator for the column covariance matrix $\mathbf{\Sigma}$ of MVND is derived under Kullback Leibler divergence loss (KLDL). Further an application of the proposed result is given in the Bayesian context of the multivariate linear model. It is illustrated that the Bayes estimators of coefficient matrix under both SEL and KLDL are identical.

Keywords: Bayes estimator, Characteristic function, Generalized matrix t-distribution, Kullback Leibler divergence loss, Matrix variate gamma distribution, Matrix variate normal distribution.

2000 Mathematics subject classification: Primary 62E05; Secondary 62E15.

*Corresponding Author

Received 01 March 2010; Accepted 25 June 2010

1. Introduction

Definition 1.1. *The random matrix $X(n\times p)$ is said to have a matrix variate normal distribution with mean $M(n\times p)$ and covariance matrix $\Omega\otimes\Sigma$ where $\Sigma(p\times p)>0$ and $\Omega(n\times n)>0$, if $vec(X)\sim N_{pn}(vec(M),\Omega\otimes\Sigma)$. We shall use the notation*
$X\sim N_{n,p}(M,\Omega\otimes\Sigma)$. The probability density function (p.d.f) of X is given by (Gupta and Nagar [11])

$$\begin{aligned} f(X) &= (2\pi)^{-np/2}\det(\Omega)^{-p/2}\det(\Sigma)^{-n/2} \\ &\quad \times\exp\left\{-\frac{1}{2}\operatorname{tr}\left[\Omega^{-1}(X-M)\Sigma^{-1}(X-M)'\right]\right\},\quad X\in\mathbb{R}^{n\times p},\quad M\in\mathbb{R}^{n\times p},\end{aligned}$$

where $\otimes$ is the Kronecker product and vec is the vectorizing operation for matrix notation.

If $X\sim N_{n,p}(M,\Omega\otimes\Sigma)$, then the chracteristic function of X is

$$\begin{aligned}\phi_X(Z) &= E\left[\exp\left(\operatorname{tr}(iZ'X)\right)\right] \\ &= \exp\left[\operatorname{tr}\left(iZ'M-\frac{1}{2}Z'\Omega Z\Sigma\right)\right],\quad Z\in\mathbb{R}^{n\times p},\end{aligned}$$

and $E(X)=M$, $COV(vec(X))=\Omega\otimes\Sigma$. (The notation $A>0$ means A is symmetric and positive definite.)
The matrix variate normal distribution (MVND) belongs to the class of matrix variate elliptically contoured distributions (Gupta and Varga [12]). In particular for $M=0$, the distribution belongs to the class of right spherical distributions, when $\Sigma=I_p$, the distribution belongs to the class of left spherical distributions when $\Omega=I_n$, and the distribution belongs to class of spherical distributions if both $\Sigma=I_p$ and $\Omega=I_n$.

The MVND has many applications, that we only focus on conditional perspective for our use. A generalized matrix t-type distribution is defined by conditioning on the MVND. More important from a practical point of view, we apply the derived results to construct the Bayes estimator and its application in multivariate linear models.
MacDonald and Newey [17] introduced the family of univariate generalized t-distributions. This family includes the normal, the power exponential and the univariate t-distributions as the special or limiting cases. Butler et al. [5] pointed out that the generalized t-distribution can be obtained as a scale mixture of the power exponential and the inverse generalized gamma distributions. This distribution has been widely used as a robust alternative to the normal distribution for modeling the errors in regression. Arellano-Valle and Bolfarine [2]

proposed a generalized multivariate t distribution family and studied the properties of the distributions included in this family. They obtained this family of distributions as a scale mixture of the normal and the inverse gamma distribution. This distribution family includes the multivariate t distribution as a special case.

Arslan [3] defined a new family of multivariate generalized distributions as a scale mixture of the multivariate power exponential distribution introduced by Gómez, et al. [10] and the inverse generalized gamma distribution with a scale parameter. Arslan [3] showed that this family of distributions belongs to the family of elliptically contoured distributions, and the multivariate normal distribution,the multivariate t distribution and the generalized multivariate t distribution introduced by Arellano-Valle and Bolfarine [2] are the special or limiting cases of the newly proposed family of multivariate generalized distributions.

Arslan [3] also showed that the univariate generalized t-distribution introduced by McDonald and Newey [17] is also a special case of this family. Thabane and Haq [19] considered the matrix generalized inverse Gaussian as the scale distribution and applied it in the Bayesian estimation of the multivariate linear model.

With these in mind, there are some other mathematical extensions by changing the scale distribution. Therefore, generalize t-models can be constracted by applying the inverse Wishart or inverse matrix variate gamma distributions.

Definition 1.2. *A random matrix Σ of order p is said to have an inverse multivariate gamma distribution with parameters α, β and Ψ denoted by $\Sigma \sim IMG_p(\alpha, \beta, \Psi)$, if its density function is given by*

$$h(\Sigma) = \frac{\det(\Psi)^{\alpha}}{\Gamma_p(\alpha)\beta^{\alpha p}} \exp\Big[-\frac{1}{\beta}\operatorname{tr}(\Psi\Sigma^{-1})\Big] \det(\Sigma)^{-\alpha-(p+1)/2},$$

where $\Sigma(p \times p) > 0$, $\Psi(p \times p) > 0$, $\alpha > (p-1)/2$, $\beta > 0$ and $\Gamma_p(\alpha)$ is the multivariate gamma function given as

$$\Gamma_p(\alpha) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma\left(\alpha - \frac{j-1}{2}\right).$$

Definition 1.3. *A random matrix Z of order p is said to have a multivariate gamma distribution with parameters α, β and Σ denoted by $Z \sim MG_p(\alpha, \beta, \Sigma)$, if its density function is given by*

$$f(Z) = \frac{\det(\Sigma)^{-\alpha}}{\beta^{p\alpha}\Gamma_p(\alpha)} \operatorname{etr}\left(-\frac{1}{\beta}\Sigma^{-1}Z\right) \det(Z)^{\alpha-(p+1)/2}, \quad Z > 0,$$

where $\alpha > (p-1)/2$, $\beta > 0$, Σ is a positive definite matrix.

2. Family of Generalized Matrix t-Distributions

In this section, the main result of the paper concerning the construction of the new family of matrix variate t-distributions is presented. This distribution will be applied in the Bayesian context.

Definition 2.1. *The random matrix $T(n \times p)$ is said to have a generalized matrix variate t-distribution (GMT) with parameters $M \in \mathbb{R}^{n\times p}$, $\Psi(p\times p) > 0$, $\Omega(n \times n) > 0$, $\alpha > (p-1)/2$, $\beta > 0$ if its p.d.f is given by*

$$\begin{aligned} f(T) &= \frac{\det(\Psi)^{-n/2}\det(\Omega)^{-p/2}\Gamma_p(\alpha+n/2)}{(2\pi/\beta)^{np/2}\Gamma_p(\alpha)} \\ &\times \det\left[I_n + \frac{\beta}{2}\Omega^{-1}(T-M)\Psi^{-1}(T-M)'\right]^{-(\alpha+n/2)}. \qquad (1) \end{aligned}$$

We shall use the notation $T \sim T_{n,p}(\alpha, \beta, M, \Omega, \Psi)$.

For $\beta = 2$ and $\alpha = \frac{n+p-1}{2}$, GMT distribution simplifies to the matrix T distribution with n degrees of freedom. (See Gupta and Nagar [11])

Theorem 2.1. *Let $X|\Sigma \sim N_{n,p}(0, \Omega \otimes \Sigma)$ and $\Sigma \sim IMG_p(\alpha, \beta, \Psi)$. Then, $X \sim T_{n,p}(\alpha, \beta, 0, \Omega, \Psi)$.*

Proof. Using conditional method, we find

$$\begin{aligned} f(X) &= \int_{\Sigma>0} g(X|\Sigma)h(\Sigma)d\,\Sigma \\ &= \frac{\det(\Psi)^{\alpha}\det(\Omega)^{-p/2}}{\Gamma_p(\alpha)(2\pi)^{np/2}\beta^{\alpha p}} \int_{\Sigma>0} \det(\Sigma)^{-(\alpha+(n+p+1)/2)} \\ &\times \exp\left[-\frac{1}{\beta}\operatorname{tr}\left((\frac{\beta}{2}X'\Omega^{-1}X+\Psi)\Sigma^{-1}\right)\right] d\,\Sigma. \qquad (2) \end{aligned}$$

Now, let $B \sim IMG_p(\alpha + n/2, \beta, \Psi_0)$. Then, from $\int_{B>0} h(B)dB = 1$ we get

$$\begin{aligned} &\int_{B>0} \det(B)^{-(\alpha+n/2)-(p+1)/2}\exp\left[-\frac{1}{\beta}\operatorname{tr}(\Psi_0 B^{-1})\right] dB \\ &= \Gamma_p\left(\alpha+\frac{n}{2}\right)\beta^{p(\alpha+n/2)}\det(\Psi_0)^{-(\alpha+n/2)}. \qquad (3) \end{aligned}$$

Substituting (3) in (2) and substituting $\Psi_0 = \frac{\beta}{2}(X'\Omega^{-1}X + \Psi)$ yields

$$f(X) = \frac{\det(\Psi)^{\alpha}\det(\Omega)^{-p/2}\Gamma_p(\alpha+n/2)}{(2\pi/\beta)^{np/2}\Gamma_p(\alpha)}\det\left(\frac{\beta}{2}X'\Omega^{-1}X+\Psi\right)^{-(\alpha+n/2)}.$$

Finally, from the following identity

$$\begin{aligned} \det\left(\Psi + \frac{\beta}{2}X'\Omega^{-1}X\right) &= \det(\Psi)\det\left(I_p + \frac{\beta}{2}\Psi^{-1}X'\Omega^{-1}X\right) \\ &= \det(\Psi)\det\left(I_n + \frac{\beta}{2}\Omega^{-1}X\Psi^{-1}X'\right), \end{aligned}$$

we obtain the result. □

3. Some properties of the GMT Family

In this section, various properties of the GMT distribution are studied using its p.d.f.

Theorem 3.1. *If $T \sim T_{n,p}(\alpha, \beta, M, \Omega, \Psi)$, then $T' \sim T_{p,n}(\alpha, \beta, M', \Psi, \Omega)$.*

Proof. By noting that

$$\det\left(I_n + \frac{\beta}{2}\Omega^{-1}(T-M)\Psi^{-1}(T-M)'\right) = \det\left(I_p + \frac{\beta}{2}\Psi^{-1}(T'-M')\Omega^{-1}(T'-M')'\right),$$

the result follows. (See page 137 Gupta and Nagar [11]) □

Theorem 3.2. *Let $T \sim T_{n,p}(\alpha, \beta, M, \Omega, \Psi)$ and $A(n \times n)$ and $B(p \times p)$ be nonsingular matrices, then $ATB \sim T_{n,p}(\alpha, \beta, AMB, A\Omega A', B'\Psi B)$.*

Proof. Transforming $W = ATB$, with the Jacobian of transformation $J(T \to W) = \det(A)^{-p}\det(B^{-n})$, from density (3) of T follows the density of W equals

$$\begin{aligned} f(W) &= \frac{\Gamma_p(\alpha + n/2)}{\Gamma_p(\alpha)(2\pi/\beta)^{np/2}} \det(\Psi)^{-n/2}\det(\Omega)^{-p/2}\det(A)^{-p}\det(B)^{-n} \\ &\quad \times \det\left(I_n + \frac{\beta}{2}\Omega^{-1}\left(A^{-1}WB^{-1} - M\right)\Psi^{-1}\left(A^{-1}WB^{-1} - M\right)'\right)^{-(\alpha+n/2)} \\ &= \frac{\Gamma_p(\alpha + n/2)}{\Gamma_p(\alpha)(2\pi/\beta)^{np/2}} \det(A\Omega A')^{-p/2}\det(B'\Psi B)^{-n/2} \\ &\quad \times \det\left(I_n + \frac{\beta}{2}(A\Omega A')^{-1}(W - AMB)(B'\Psi B)^{-1}(W - AMB)'\right)^{-(\alpha+n/2)}, \end{aligned}$$

where $W \in \mathbb{R}^{n\times p}$, and, hence the result. □

Theorem 3.3. *Let $T \sim T_{n,p}(\alpha, \beta, M, \Omega, \Psi)$, then the characteristic function of T is given by*

$$\phi_T(Z) = \frac{\exp[\mathrm{tr}(iZ'M)]\det(\Psi)^{\alpha}}{\Gamma_p(\alpha)(2\beta)^{\alpha p}} \det(Z'\Omega Z)^{\alpha} B_{\alpha}(\frac{1}{2\beta}Z'\Omega Z\Psi),$$

Further, $E(T) = M$ and, $COV(T) = \frac{2(\Omega\otimes\Psi)}{\beta(2\alpha-n-1)}$.

Proof.

$$\begin{aligned} \phi_T(Z) &= E[\exp(\mathrm{tr}(iZ'T))], \quad Z \in \mathbb{R}^{n\times p} \\ &= E[E(\exp(\mathrm{tr}(iZ'T))|\Sigma)] \\ &= E[\exp(\mathrm{tr}(iZ'M - \frac{1}{2}Z'\Omega Z\Sigma))] \\ &= e^{\mathrm{tr}(iZ'M)}\int_{\Sigma>0}\exp[-\frac{1}{2}\mathrm{tr}(Z'\Omega Z\Sigma)]g(\Sigma)d\,\Sigma, \end{aligned}$$

where $\Sigma \sim IMG_p(\alpha, \beta, \Psi)$, then

$$\begin{aligned}
\phi_T(Z) &= \frac{\exp[\mathrm{tr}(iZ'M)]\det(\Psi)^{\alpha}}{\Gamma_p(\alpha)\beta^{\alpha p}} \\
&\quad \times \int_{\Sigma>0} \det(\Sigma)^{-\alpha-(p+1)/2} \exp[-\,\mathrm{tr}(\frac{1}{\beta}\Psi\Sigma^{-1} + \frac{1}{2}Z'\Omega Z\Sigma)]d\,\Sigma \\
&= \frac{\exp[\mathrm{tr}(iZ'M)]\det(\Psi)^{\alpha}}{\Gamma_p(\alpha)(2\beta)^{\alpha p}} \det(Z'\Omega Z)^{\alpha} B_{\alpha}(\frac{1}{2\beta}Z'\Omega Z\Psi),
\end{aligned}$$

where B_δ is the type two Bessel function of Herz of matrix argument. B_δ is defined as

$$B_\delta(WZ) = \det(W)^{-\delta}\int_{S>0} \exp\left(tr(-SW - S^{-1}Z)\right)\det(S)^{-\delta-\frac{1}{2}(p+1)}dS,$$

(See p39 Gupta and Nagar [11] and Herz [15].)
Using conditional expectation we have

$$\begin{aligned}
E(T) &= E(E(T|\Sigma)) \\
&= E(M) = M, \\
Cov(T) &= E(Cov(T|\Sigma)) \\
&= \frac{2E(\Omega\otimes\Sigma)}{\beta(2\alpha-n-1)} \\
&= \frac{2(\Omega\otimes\Psi)}{\beta(2\alpha-n-1)}.
\end{aligned}$$

□

In the following section another point of view is considered concerning the result obtained in Theorem 2.1. This result is an extension to the work of Dickey [8].

Theorem 3.4. *Let $X \sim N_{n,p}(0, I_n \otimes \Omega)$, independent of $S \sim MG_n(\alpha, \beta, \Lambda^{-1})$ define*

$$T = (S^{-1/2})'X + M,$$

where $M(n \times p)$ is a constant matrix and $S^{1/2}(S^{1/2})' = S$. Then, $T \sim T_{n,p}(\alpha, \beta, M, \Lambda, \Omega)$.

Proof. The joint density of X and S is given by

$$\begin{aligned}
f(X, S) &= \frac{(2\pi)^{-np/2}\det(\Omega)^{-n/2}\det(\Lambda)^{\alpha}}{\beta^{n\alpha}\Gamma_n(\alpha)} \det(S)^{\alpha-(n+1)/2} \\
&\quad \times \exp\left[-\,\mathrm{tr}\left(\frac{1}{\beta}\Lambda S + \frac{1}{2}X\Omega^{-1}X'\right)\right], \quad S > 0, \quad X \in \mathbb{R}^{n\times p}.
\end{aligned}$$

Now, let $T = (S^{-1/2})'X + M$. The Jacobian of transformation is $J(X \to T) = \det(S)^{p/2}$. Substituting for X in terms of T in the joint density of X and S, and multiplying the resulting expression by$J(X \to T)$, we get the joint p.d.f of T and S as

$$\begin{aligned} f(T,S) &= \frac{(2\pi)^{-np/2}\det(\Omega)^{-n/2}\det(\Lambda)^{\alpha}}{\beta^{n\alpha}\Gamma_n(\alpha)}\det(S)^{\alpha-(n+1)/2+p/2} \\ &\quad \times \exp\left[-\operatorname{tr}\left(\frac{1}{\beta}\Lambda+\frac{1}{2}(T-M)\Omega^{-1}(T-M)'\right)S\right], \quad S>0, \quad T\in\mathbb{R}^{n\times p} \end{aligned}$$

Now integrating out S using multivariate gamma integral as (Anderson [1])

$$\int_{Z>0}\det(Z)^{\alpha-(p+1)/2}\exp[-\operatorname{tr}(\Lambda Z)]dZ = \det(\Lambda)^{-\alpha}\Gamma_p(\alpha).$$

Then the density of T is obtained as

$$\begin{aligned} f(T) &= \frac{(2\pi)^{-np/2}\det(\Omega)^{-n/2}\det(\Lambda)^{\alpha}\Gamma_n(\alpha+p/2)}{\beta^{n\alpha}\Gamma_n(\alpha)} \\ &\quad \times \det\left(\frac{1}{\beta}\left(\Lambda+\frac{\beta}{2}(T-M)\Omega^{-1}(T-M)'\right)\right)^{-(\alpha+p/2)} \\ &= \frac{\det(\Omega)^{-n/2}\det(\Lambda)^{-p/2}\Gamma_n(\alpha+p/2)}{(2\pi/\beta)^{np/2}\Gamma_n(\alpha)} \\ &\quad \times \det\left(I_n+\frac{\beta}{2}\Lambda^{-1}(T-M)\Omega^{-1}(T-M)'\right)^{-(\alpha+p/2)}. \end{aligned}$$

□

4. Bayes Estimation

Over the years, considerable research has been done in the field of the Baysian estimation. For example, Bekker and Roux [4] considered the Bayesian analysis of the multivariate normal distribution when its covariance matrix has a Wishart prior under quadratic loss. In the same spirit, Haff [13, 14], Dey and Srinivasan [7] and Towhidi and Behboodian [20] applied entropy and extended reflected normal loss functions in the Bayesian estimation of covariance matrix of the multivariate normal distribution respectively. Thabane and Haq [19] derived the Bayes estimator in the matrix variate normal distribution under square error loss (SEL) function. They applied their results in Bayesian estimation of the multivariate linear models. Díaz-García and Gutiérrez [9] also considered the distribution of a random singular matrix as a prior to the Bayesian inference of the multivariate linear models. For application purposes, in this section the Bayes estimator of Σ based on the conditional property is derived. In this regard, we consider Kullback Leibler divergence loss (KLDL) as the measurement. First we state a result due to Das and Dey [6].

Lemma 4.1. *Suppose A is an estimator for unknown parameters matrix Σ , where $\pi(A|D)$ and $\pi(\Sigma|D)$ are the corresponding posterior probability density function over $\mathcal{R}^p$ respectively, where D indicates Data. Now the posterior* *expected loss of A, when the posterior distribution is* $\pi(\Sigma|D)$, *is*

$$\rho(\Sigma, A) = \int_{\Sigma>0} L(\Sigma, A) dF^{\pi(\Sigma|D)}(\Sigma),$$

where the loss function is $\log\left(\frac{\pi(A|D)}{\pi(\Sigma|D)}\right)$. *Then the Bayes rule is*

$$\delta^{\pi}(D) = argmax_{\Sigma>0}\pi(\Sigma|D),$$

which is the mode of posterior distribution of $\pi(\Sigma|D)$.

Note that the posterior expected loss function in Lemma 4.1,

$$\rho(\Sigma, A) = E\left[\log\left(\frac{\pi(A|D)}{\pi(\Sigma|D)}\right)\right],$$

can be interpreted as the Kullback Leibler divergence of the posterior distribution evaluated under action A from the true posterior distribution of unknown parameters Σ. Therefore the posterior expected loss or Kullback Leibler divergence is minimum, if we choose our action as posterior mode.

Lemma 4.2. *Let* $X|\Sigma \sim N_{n,p}(0, \Omega \otimes \Sigma)$. *Further* Σ *has prior distribution as* $IMG_p(\alpha, \beta, \Psi)$. *Then the posterior distribution of* Σ *is*

$$\begin{aligned}\pi(\Sigma|X) &= \frac{\det(\Psi)^{\alpha+n/2}\det(\Sigma)^{-(\alpha+n/2+(p+1)/2)}}{\Gamma_p(\alpha+n/2)\beta^{p\alpha+np/2}} \\ &\times \det\left(I_n + \frac{\beta}{2}\Omega^{-1}X\Psi^{-1}X'\right)^{\alpha+n/2} \exp\left(-\operatorname{tr}(\frac{1}{2}\Omega^{-1}X\Sigma^{-1}X') - tr(\frac{1}{\beta}\Psi\Sigma^{-1})\right).\end{aligned}$$

Proof. By definition

$$\pi(\Sigma|X) = \frac{f(X|\Sigma)\pi(\Sigma)}{m(X)}.$$

By applying Theorem 2.1 we obtain the underlying result. □

Theorem 4.1. *Suppose* $X|\Sigma \sim N_{n,p}(0, \Omega \otimes \Sigma)$.*Further* Σ *has the prior distribution as* $IMG_p(\alpha, \beta, \Psi)$.*Under KLDL function,the Bayes estimation of* Σ *is given by*

$$\widehat{\Sigma} = [\alpha + n/2 + (p+1)/2]^{-1}\left(\frac{1}{2}X'\Omega^{-1}X + \frac{1}{\beta}\Psi\right).$$

Proof. Taking logarithm on both sides of $\pi(\Sigma|X)$ in Lemma 4.2 yields

$$\begin{aligned}\log[\pi(\Sigma|X)] &= constant - (\alpha + n/2 + (p+1)/2)\log\det(\Sigma) \\ &- \frac{1}{2}\operatorname{tr}(\Omega^{-1}X\Sigma^{-1}X') - \frac{1}{\beta}\operatorname{tr}(\Psi\Sigma^{-1}).\end{aligned}$$

Now differentiating with respect to Σ and equating the derivative to 0, gives (Seber [18] and Mardia et al. [16])

$$\begin{aligned}\frac{\partial \log \pi(\Sigma|X)}{\partial \Sigma} &= - (\alpha + n/2 + (p+1)/2)[2\Sigma^{-1} - diag\Sigma^{-1}] \\ &+ \frac{1}{2}[2\Sigma^{-1}X'\Omega^{-1}X\Sigma^{-1} - diag(\Sigma^{-1}X'\Omega^{-1}X\Sigma^{-1})] \\ &+ \frac{1}{\beta}[2\Sigma^{-1}\Psi\Sigma^{-1} - diag(\Sigma^{-1}\Psi\Sigma^{-1})] \\ &= 2M - diagM = 0,\end{aligned}$$

where

$$M = -(\alpha + n/2 + (p+1)/2)\Sigma^{-1} + \frac{1}{2}\Sigma^{-1}X'\Omega^{-1}X\Sigma^{-1} + \frac{1}{\beta}\Sigma^{-1}\Psi\Sigma^{-1}.$$

Then

$$\Sigma^{-1}[-(\alpha + n/2 + (p+1)/2)\Sigma + \frac{1}{2}X'\Omega^{-1}X + \frac{1}{\beta}\Psi]\Sigma^{-1} = 0,$$

which completes the proof. □

5. Application to the multivariate linear model

In this section, following Thabane and Haq [19] we consider the Bayesian application of the proposed results in the multivariate linear model. As a precise setup consider the multivariate linear model

$$Y = XB + E,$$

where Y is a $n \times p$ matrix of observed responses, X is a $n \times m$ matrix of fixed elements with $rank(X) = m$, B is a $m \times p$ matrix of unknown regression parameters, and E is a matrix of errors with $E \sim N_{n\times p}(\mathbf{0}, I_n \otimes \Sigma)$.
The likelihood function is given by

$$f(Y|B,\Sigma) = (2\pi)^{-np/2}\det(\Sigma)^{-n/2}\exp\left[-\frac{1}{2}tr(Y - XB)'(Y - XB)\Sigma^{-1}\right].$$

We assume that our prior information about the parameter space of (B, Σ) is summarized by $B|\Sigma \sim N_{m\times p}(B_0, (X'X)^{-1} \otimes \Sigma)$ and $\Sigma \sim IMG_p(\alpha, \beta, \Psi)$. Thus The conjugate prior densities are

$$f(B|\Sigma) = (2\pi)^{-mp/2}\det(X'X)^{p/2}\det(\Sigma)^{-m/2}\exp\left[-\frac{1}{2}tr(B - B_0)'X'X(B - B_0)\Sigma^{-1}\right],$$

and

$$f(\Sigma) = \frac{\det(\Psi)^{\alpha}}{\Gamma_p(\alpha)\beta^{\alpha p}}\exp\Big[-\frac{1}{\beta}\,\mathrm{tr}(\Psi\Sigma^{-1})\Big]\det(\Sigma)^{-\alpha-(p+1)/2},$$

where $\Sigma(p \times p) > 0$ and $\Psi(p \times p) > 0$.
Combining the likelihood function with the prior, then after a few matrix algebraic

steps, the posterior density is

$$f(B,\Sigma|Y) \propto \det(\Sigma)^{-(\alpha+\frac{m+n}{2})-\frac{p+1}{2}} \times \exp\left[\frac{-1}{2}tr\left[(B-\widehat{B}_*)'(2X'X)(B-\widehat{B}_*)+S_*\right]\Sigma^{-1}\right], \tag{4}$$

where

$$\begin{aligned}
\widehat{B}_* &= \frac{B_0+\widehat{B}}{2},\\
S &= (Y-X\widehat{B})'(Y-X\widehat{B}),\\
\widehat{B} &= (X'X)^{-1}X'Y,\\
S_* &= \frac{2}{\beta}\Psi+S+(\widehat{B}-B_0)'(2X'X)^{-1}(\widehat{B}-B_0).
\end{aligned}$$

The marginal posterior density of B is obtained by integrating Eq.(4) with respect to Σ

$$\begin{aligned}
f(B|Y) &= \int_{\Sigma>0} f(B,\Sigma|Y)d\Sigma\\
&\propto \int_{\Sigma>0} \det(\Sigma)^{-(\alpha+\frac{m+n}{2})-\frac{p+1}{2}}\\
&\times \exp\left[\frac{-1}{2}tr\left[(B-\widehat{B}_*)'(2X'X)(B-\widehat{B}_*)+S_*\right]\Sigma^{-1}\right]d\Sigma.
\end{aligned} \tag{5}$$

we note that the posterior density (5) can be write as

$$f(B|Y) = \int_{\Sigma>0} f(B|\Sigma,Y)f(\Sigma|Y)d\Sigma, \tag{6}$$

where

$$\begin{aligned}
f(B|\Sigma,Y) &= (2\pi)^{\frac{-mp}{2}}\det(2X'X)^{\frac{p}{2}}\det(\Sigma)^{\frac{-m}{2}}\\
&\times \exp\left[-\frac{1}{2}tr(B-\widehat{B}_*)'(2X'X)(B-\widehat{B}_*)\Sigma^{-1}\right],
\end{aligned} \tag{7}$$

$$f(\Sigma|Y) = \frac{\det(S_*)^{-(\alpha+\frac{n}{2})}\det(\Sigma)^{-(\alpha+\frac{n}{2})-\frac{p+1}{2}}}{\Gamma_p(\alpha+\frac{n}{2})\beta^{p(\alpha+n/2)}}\exp\left[-\frac{1}{\beta}tr(S_*\Sigma^{-1})\right]. \tag{8}$$

Substituting (7) and (8) in (6) and using theorem 2.1 we get

$$B|Y \sim T_{m,p}\left(\alpha+n/2,\beta,\widehat{B_0},S_*,(2X'X)^{-1}\right).$$

From Theorem 3.3 $E(B|Y)=\widehat{B_*}$ and the mode of $B|Y$ is $\widehat{B_*}$, since having a symmetric distribution. Thus under SEL and KLDL the Bayesian estimate of B is identical and equal to $\widehat{B_*}$.

Acknowledgments. The authors would like to thanks two anonymous referees for constructive and helpful comments which significantly improved the context of the paper.

References

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. Third edition, Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, N.J., 2003.

[2] Reinaldo B. Arellano-Valle and Heleno Bolfarine, On some characterizations of the t-distribution, *Statist. Probab. Lett.*, **25**(1) (1995), 79–85.

[3] Olcay Arslan, Family of multivariate generalized t distributions, *J. Mult. Anal.*, **89**(2) (2004), 329–337.

[4] A. Bekker, J. J. J. Roux, Baysian multivariate normal analysis with a Wishart prior. *Comm. Statist. Theo. Meth.*, **24**(10) (1995), 2485-2497.

[5] R. J. Butler, J. B. McDonald, R. D. Nelson and S. B. White, Robust and partially adaptive estimation of regression models, *Rev. Econom. Statist.*, **72** (1990), 321–327.

[6] S. Das and D. K. Dey, On Bayesian inference for generalized multivariate gamma distribution, *Statist. Prob. Lett.*, (2010), Doi:10.1016/j.spl.2010.05.018.

[7] D. K. Dey, C. Srinivasan, Estimation of the inverse covariance matrix under Stein's loss, *Ann. statist.*, **13**(4) (1985), 1581-1591.

[8] J. M. Dickey Matrixvariate generalizations of the multivariate t and the inverted multivariate t distribution, *Ann. Math. Statist.*, **38** (1967), 511-518.

[9] J. A. Díaz-García, J. R. Gutiérrez, Distribution of the generalized inverse of a random matrix and its applications, *J. Statist. Plann. Inf.*, **136** (2006), 183-192.

[10] E. Gómez, M. A. Gómez-Villegas and J. M. Marín, A multivariate generalization of the power exponential family of distributions, *Comm. Statist. Theo. Meth.*, **27**(3) (1998), 589–600.

[11] A. K. Gupta and D. K. Nagar, *Matrix Variate Distribution*, Chapman and Hall/CRC,Boca Raton, FL, 2000.

[12] A. K. Gupta and T. varga, *Elliptically Contoured Model in Statistics*, Kluwer Academic Press, London, 1993.

[13] L. R. Haff, Estimation of the inverse covariance matrix mixtures of the inverse Wishart matrix and an identity,*Ann. Statist.*, **7**, (1979), 1264-1276.

[14] L. R. Haff, Empirical Bayes estimation of the multivariate covariance matrix, *Ann. Statist.*, **8**,(1980), 586-597.

[15] Carl S. Herz, Bessel functions of matrix argument, *Ann. Math.*, **61**(2) (1955), 474-523.

[16] Kantilal V. Mardia, John T. Kent and John M. Bibby, *Multivariate analysis, Probability and Mathematical Statistics: A Series of Monographs and Textbooks*, Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York-Toronto, Ont., 1979.

[17] James B. McDonald and Whitney K. Newey, Partially adaptive estimation of regression models via the generalized t distribution, *Econometric Theory*, **4**(3) (1988), 428–457.

[18] George A. F. Seber, *A Matrix Handbook For Statisticians* , John Wiley, 2007.

[19] L. Thabane, M. Safiul. Haq, On the matrix-variate generalized hyperbolic distribution and its Bayesian applications, *Statistics*, **38**(6) (2004), 511-526.

[20] M. Towhidi, J.Behboodian, Bayesian multivariate normal analysis under the extended reflected normal loss function, *Statistics*, **35**(3) (2001), 215-223.