Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Multivariate Analysis 97 (2006) 785–795
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

The matrix-t distribution and its applications in predictive inference

B.M. Golam Kibria*

Department of Statistics, Florida International University, University Park, Miami, FL 33199, USA

Received 26 August 2004
Available online 16 September 2005

Abstract

The predictive distributions of the future responses and regression matrix under the multivariate elliptically contoured distributions are derived using structural approach. The predictive distributions are obtained as matrix-t which are identical to those obtained under matrix normal and matrix-t distributions. This gives inference robustness with respect to departures from the reference case of independent sampling from the matrix normal or dependent but uncorrelated sampling from matrix-t distributions. Some successful applications of matrix-t distribution in the field of spatial prediction have been addressed.

AMS 1991 subject classification: primary 62F15; secondary 62J05; 62E15

Keywords: Elliptically contoured distribution; Matrix-t; Multivariate linear model; Predictive distribution; Robustness; Spatial prediction; Tolerance region

1. Introduction

The predictive inference for the multivariate Gaussian regression models has been considered by various researchers: Geisser [10] and Guttman and Hougaard [12] considered the classical approach, Zellner and Chetty [29] and Kibria et al. [17] considered the Bayesian method while Fraser and Haq [8], and Haq [13] considered the structural relation of the model approach to mention a few. The assumption of normality and independency for the

* Fax: +305 348 6895.
E-mail address: kibriag@fiu.edu.

doi:10.1016/j.jmva.2005.08.001

error variables may not be appropriate in many practical situations, specially when the underlying distributions have heavier tails. For such cases, multivariate t-errors with linear models have been emphasized by Zellner [28] and Sutradhar and Ali [27] among others. In the case of multivariate linear model, matrix-t errors have been considered by Kibria and Haq [16] and Khan [14] among others. The predictive distribution remains the same by a change in the error distribution from normal to multivariate t [28] or from normal to elliptical distribution [9].

The literature on predictive distribution for the future regression matrix is limited. Haq [13] used the structural relation of the model to derive the predictive distribution for the future response and regression matrices under the matrix normal error assumption. For both cases, he obtained predictive distributions as matrix-t with appropriate degrees of freedom. Kibria and Haq [16] considered the predictive distribution for future responses under the matrix-t errors and obtained the predictive distribution as a matrix-t with appropriate degrees of freedom. Khan [14] considered the structural relation to derive the predictive distribution of regression matrix under the matrix-t error and obtained the predictive distribution as a matrix-t with appropriate degrees of freedom. Therefore, the distribution of future regression matrix is unaffected by a change in the error distribution from matrix normal to matrix-t distribution. This invariance principle suggests that the predictive distribution would be invariant to a wide class of error distributions, namely multivariate elliptically contoured distribution.

The class of elliptically contoured distributions includes various distributions: the multivariate normal, matrix-t, multivariate Student's t and multivariate Cauchy (see [23]). The class of mixture of normal distributions is a subclass of the class of elliptical distributions as well as the class of spherically symmetric distribution [5]. Elliptically contoured distributions have been discussed extensively for traditional multivariate regression model by Anderson and Fang [1], Fang and Li [6], Fang and Zhang [7], Gupta and Varga [11] and Kubokawa and Srivastava [19]. These distributions have also been considered by Chib et al. [4], Kibria and Haq [15], Ng [23,24] and Kim and Mallick [18] in the context of predictive inference for the future responses but not for the future regression matrix.

In this paper, a very general assumption is employed, namely that responses have a multivariate elliptically contoured distribution. It has been shown that the prediction distribution of future response and regression matrix are obtained as matrix-t distributions. Therefore, the assumption of normality as well as matrix-t are robust to deviation in the direction of elliptical distribution as far as inference about the future regression matrix and prediction are concerned. Since the predictive distributions are matrix-t, this paper will address some applications of matrix-t distribution. A random matrix $X_{n\times m}$ is said to have a matrix-t distribution (in the notation of Box and Tiao [2, p. 442]) with v degrees of freedom if its joint probability density function is expressible as

$$f(X) = C|I_n + v^{-1}\{A^{-1}(X-\mu)\}\{(X-\mu)B^{-1}\}'|^{-\frac{v+n+m-1}{2}}, \tag{1.1}$$

where

$$C = \frac{|A|^{-m/2}|B|^{-n/2}\Gamma_{n+m}((v+n+m-1)/2)}{(v\pi^2)^{nm/2}\Gamma_n((v+n-1)/2)\Gamma_m((v+m-1)/2)}$$

is the normalizing constant. Also $\Gamma_p(\lambda) = \pi^{p(p-1)/4}\Gamma(\lambda)\Gamma(\lambda - \frac{1}{2}) \ldots \Gamma(\lambda - 1/2p + \frac{1}{2})$ is the multivariate gamma function and v is the shape parameter. The mean and variance of the matrix X are, respectively, $E(X) = \mu$ and $V(X) = \frac{1}{v-2} A \otimes B$, where $\otimes$ is the Kronecker product between two matrices. We write $X \sim t_{m\times n}(\mu, A, B, v)$, where $A_{n\times n}$ and $B_{m\times m}$ are positive definite matrices. The matrix-t distribution and some of its properties are discussed in Loschi et al. [22] and Press [25] among others. The organization of this paper is as follows. The predictive distributions for future response and regression matrices are derived in Section 2. Some applications of matrix-t distribution are discussed in Section 3. Finally, some concluding remarks are given in Section 4.

2. Predictive distributions

Ng [23] derived the predictive distribution of future response matrix for a multivariate linear model with an elliptically contoured error distribution. He has considered both classical and Bayesian with improper prior for the derivation of the predictive distribution. This section will discuss about the derivation of the predictive distributions of future response and regression matrices for the model (2.2) using the structural relation of the model approach developed by Fraser and Haq [8].

2.1. The model and some preliminaries

We consider the following multivariate linear model

$$Y = \beta X + \Gamma E, \tag{2.2}$$

where Y is an $m \times n$ matrix of observed responses, β is an $m \times p$ matrix of regression parameters, X is an $p \times n$ $(n \geqslant p)$ known regression matrix, Γ is an $m \times m$ matrix of scale parameter with $|\Gamma| > 0$ and E is an $m \times n$ random error matrix. We assume that E has a spherically contoured distribution with the probability density function

$$f(E) \propto g\{\mathrm{tr}(EE')\}. \tag{2.3}$$

Then the response matrix Y has an elliptically contoured distribution with *pdf* as

$$f(Y|\beta, \Sigma) \propto |\Sigma|^{-\frac{n}{2}} g\{\mathrm{tr}\,\Sigma^{-1}(Y - \beta X)(Y - \beta X)'\}, \tag{2.4}$$

where $\Gamma\Gamma' = \Sigma$. This is the form given in Anderson and Fang [1], where $g\{\cdot\}$ is a non-negative function over $m \times m$ positive definite matrices such that $f(Y|\beta, \Sigma)$ is a density function. Here M' denotes the transpose of the matrix M and $\mathrm{tr}(M)$ denotes the trace of the matrix M.

We define

$$B_E = EX'(XX')^{-1} \quad \text{and} \quad S_E = (E - B_E X)(E - B_E X)', \tag{2.5}$$

as the regression matrix of E on X and the sum of squares and product (SSP) matrix, respectively. Consider C_E be a non-singular matrix such that the error SSP matrix, S_E

can be expressed as

$$C_E C_E' = S_E,$$

and the standardized residual matrix is

$$W_E = C_E^{-1}(E - B_E X). \tag{2.6}$$

From Eq. (2.6), we have the following relationship:

$$E = B_E X + C_E W_E. \tag{2.7}$$

It is clear that $W_E W_E' = I_m$, we have

$$EE' = B_E XX' B_E' + C_E C_E'. \tag{2.8}$$

Similarly we can define the quantities in (2.5)–(2.8) in terms of future errors, observed and future responses. The following two subsections are devoted to deriving the predictive distribution of future response and regression matrices, respectively.

2.2. Predictive distribution of future response matrix

Consider a set of n_{f} future responses from the multivariate linear model defined in (2.2) as

$$Y_{\mathrm{f}} = \beta X_{\mathrm{f}} + \Gamma E_{\mathrm{f}}, \tag{2.9}$$

where Y_{f} and E_{f} are the $m \times n_{\mathrm{f}}$ matrices of future responses and errors, respectively, and X_{f} is an $p \times n_{\mathrm{f}}$ ($n_{\mathrm{f}} \geqslant p$) future regression matrix. It is assumed that E_{f} has the same realization as E, then the joint distribution of E and E_{f} can be written as

$$f(E, E_{\mathrm{f}}) \propto g\{\mathrm{tr}(EE' + E_{\mathrm{f}} E_{\mathrm{f}}')\}. \tag{2.10}$$

Following Fraser and Ng [9], the relationship between volume elements of E in terms of B_E, S_E and W_E is given by

$$dE \propto |XX'|^{\frac{m}{2}} |S_E|^{\frac{n-p-m-1}{2}} dB_E \, dS_E \, dW_E. \tag{2.11}$$

Now considering (2.7) and (2.10) and taking into account the Jacobian (2.11), the joint density of function of B_E, S_E and E_{f} is obtained as

$$f(B_E, S_E, E_{\mathrm{f}}|E, X) \propto |S_E|^{\frac{n-p-m-1}{2}} g\{\mathrm{tr}(S_E + B_E XX' B_E + E_{\mathrm{f}} E_{\mathrm{f}}')\}. \tag{2.12}$$

Making the following transformation:

$$R = S_E^{-\frac{1}{2}}(E_{\mathrm{f}} - B_E X_{\mathrm{f}}),$$

the Jacobian of the transformation is $|S_E|^{\frac{n_{\mathrm{f}}}{2}}$, thus we have the joint density of B_E, S_E, R as

$$f(B_E, S_E, R|E, X, X_{\mathrm{f}}) \propto |S_E|^{\frac{n+n_{\mathrm{f}}-p-m-1}{2}} g\{\mathrm{tr}(S_E + B_E XX' B_E + (S_E^{\frac{1}{2}} R + B_E X_{\mathrm{f}})(S_E^{\frac{1}{2}} R + B_E X_{\mathrm{f}})')\}. \tag{2.13}$$

The matrix expression under trace in (2.13) can be expressed as

$$\operatorname{tr}((I_m + RHR')S_E + (B_E + S_E^{\frac{1}{2}} R X_{\mathrm{f}}' A^{-1}) A (B_E + S_E^{\frac{1}{2}} R X_{\mathrm{f}}' A^{-1})'),$$

where $H = I_{n_{\mathrm{f}}} - X_{\mathrm{f}}' A^{-1} X_{\mathrm{f}}$ and $A = XX' + X_{\mathrm{f}} X_{\mathrm{f}}'$. Now assume $I_m + RHR'$ is positive definite and Q is a non-singular matrix such that $Q'Q = I_m + RHR'$, and make the following transformation:

$$K = QS_E Q',$$
$$Z = B_E + S_E^{\frac{1}{2}} R X_{\mathrm{f}}' A^{-1}.$$

Then the Jacobian of the transformation $J(B_E, S_E, R) \to (Z, K, R)$ equals to $|Q|^{-(m+1)}$, and we obtain the joint density function of Z, K and R is as follows:

$$p(Z, K, R|E, X, X_{\mathrm{f}}) \propto |I_m + RHR'|^{\frac{n+n_{\mathrm{f}}-p}{2}} |K|^{\frac{n+n_{\mathrm{f}}-m-p-1}{2}} g\{\operatorname{tr}(K) + \operatorname{tr}(ZAZ')\}. \quad (2.14)$$

Integrating (2.14) with respect to Z and K yields the density function of R as

$$\int_K \int_Z p(R, K, Z|E, X, X_{\mathrm{f}})\, dZ\, dK \propto |I_m + RHR'|^{\frac{n+n_{\mathrm{f}}-p}{2}}.$$

It can be shown that,

$$R = S_E^{-\frac{1}{2}}(E_{\mathrm{f}} - B_E X_{\mathrm{f}}) = S_Y^{-\frac{1}{2}}(Y_{\mathrm{f}} - B_Y X_{\mathrm{f}}).$$

Finally, the *pdf* of Y_{f} for given data is obtained as

$$p(Y_{\mathrm{f}}|Y, X, X_{\mathrm{f}}) \propto |I_m + S_Y^{-1}(Y_{\mathrm{f}} - B_Y X_{\mathrm{f}})(I_{n_{\mathrm{f}}} - X_{\mathrm{f}}' A^{-1} X_{\mathrm{f}})(Y_{\mathrm{f}} - B_Y X_{\mathrm{f}})'|^{\frac{n+n_{\mathrm{f}}-p}{2}}. \quad (2.15)$$

From (1.1) and (2.15), it follows that Y_{f} has a matrix-t density. Thus the predictive distribution of the future responses for given data is an $m \times n_{\mathrm{f}}$ dimensional matrix-t distribution with $(n - p - m + 1)$ degrees of freedom. That is

$$Y_{\mathrm{f}} \sim t_{m \times n_{\mathrm{f}}}(B_Y X_{\mathrm{f}}, (I_{n_{\mathrm{f}}} - X_{\mathrm{f}}' A^{-1} X_{\mathrm{f}})^{-1}, S_Y, n - p - m + 1).$$

This result coincides with that of Haq [13], where he considered the normal error and after little modification with that of Kibria and Haq [16], where they considered matrix-t error. Thus the predictive distribution is unaffected by departures from normality or dependent but uncorrelated assumptions to elliptically contoured distribution. This result also coincides with that of Ng [23], where he obtained the predictive distribution of future responses for elliptically contoured distribution under both classical and Bayesian approaches with improper prior distribution. Thus, it can be concluded that the predictive distribution for multivariate linear model with elliptically contoured distribution under the improper Bayesian, classical and structural approaches are the same.

2.3. Predictive distribution of future regression matrix

Haq [13] and Khan [14] derived the predictive distributions of future regression matrix for a multivariate linear model using the matrix normal and matrix-t error, respectively.

Both considered the structural relation of the model to derive the predictive distributions and obtained predictive distributions as matrix-t with appropriate degrees of freedom. This section will consider the derivation of predictive distributions for future regression matrix and sum of squares and product matrix using structural relation of the model when the error of the model have elliptically contoured distribution. Following Fraser and Ng [9] and Khan [14], the joint density function of error statistics B_E, S_E, B_{E_f} and S_{E_f} is obtained as

$$p(B_E, S_E, B_{E_\mathrm{f}}, S_{E_\mathrm{f}}|E, X, X_\mathrm{f}) \propto |S_E|^{\frac{n-m-p-1}{2}}|S_{E_\mathrm{f}}|^{\frac{n_\mathrm{f}-m-p-1}{2}} g\{\mathrm{tr}(B_E XX'B_E' + S_E + B_{E_\mathrm{f}} X_\mathrm{f} X_\mathrm{f}' B_{E_\mathrm{f}}' + S_{E_\mathrm{f}})\}. \tag{2.16}$$

The structural relation of the model (2.2) yields

$$B_E = \Sigma^{-\frac{1}{2}}(B_Y - \beta) \quad \text{and} \quad S_E = \Sigma^{-1} S_Y,$$

where B_Y is the regression matrix of Y on X, and $S_Y = (Y - B_Y)(Y - B_Y)'$ is the Wishart matrix. The Jacobian of the transformation $J\{[B_E, S_E] \to [\beta, \Sigma]\}$ is equal to $|S_Y|^{\frac{m+1}{2}}|\Sigma|^{-(\frac{p}{2}+m+1)}$. Then the joint density of β, Σ, B_{E_f}, and S_{E_f} is obtained as

$$p(\beta, \Sigma, B_{E_\mathrm{f}}, S_{E_\mathrm{f}}|E, X, X_\mathrm{f}) \propto |S_{E_\mathrm{f}}|^{\frac{n_\mathrm{f}-m-p-1}{2}}|\Sigma|^{-\frac{n+m+1}{2}} g\{\mathrm{tr}\,\Sigma^{-1}((B-\beta)XX'(B-\beta)' + S + B_{E_\mathrm{f}} X_\mathrm{f} X_\mathrm{f}' B_{E_\mathrm{f}}' + S_{E_\mathrm{f}})\},$$

where $B_Y = B = (b_1, b_2, \ldots, b_m)'$ and $S_Y = S$ for notational convenience. Similarly, the structural relation of the model (2.9) yields

$$B_{E_\mathrm{f}} = \Sigma^{-\frac{1}{2}}(B_{Y_\mathrm{f}} - \beta) \quad \text{and} \quad S_{E_\mathrm{f}} = \Sigma^{-1} S_{Y_\mathrm{f}},$$

where $B_{Y_\mathrm{f}} = (b_{\mathrm{f}1}, b_{\mathrm{f}2}, \ldots, b_{\mathrm{f}m})'$ is the regression matrix of Y_f on X_f, and S_{Y_f} is the Wishart matrix for the future responses. The Jacobian of the transformation $J\{[B_{E_\mathrm{f}}, S_{E_\mathrm{f}}] \to [B_\mathrm{f}, S_\mathrm{f}]\}$ is equal to $|\Sigma|^{-\frac{p+m+1}{2}}$. Then the joint density function of β, Σ, B_f, and S_f is obtained as

$$p(\beta, \Sigma, B_\mathrm{f}, S_\mathrm{f}|Y, X, X_\mathrm{f}) \propto |S_\mathrm{f}|^{\frac{n_\mathrm{f}-m-p-1}{2}}|\Sigma|^{-\frac{n+n_\mathrm{f}+m+1}{2}} g\{\mathrm{tr}(\Sigma^{-1}[(B-\beta)XX'(B-\beta)' + S + (B_\mathrm{f}-\beta)X_\mathrm{f}X_\mathrm{f}'(B_\mathrm{f}-\beta)' + S_\mathrm{f}])\}, \tag{2.17}$$

where $B_{Y_\mathrm{f}} = B_\mathrm{f}$ and $S_{Y_\mathrm{f}} = S_\mathrm{f}$.

The joint density function of β, B_f and S_f is obtained from (2.17) as

$$p(\beta, B_\mathrm{f}, S_\mathrm{f}|Y, X, X_\mathrm{f}) \propto |S_\mathrm{f}|^{\frac{n_\mathrm{f}-m-p-1}{2}} \int_\Sigma |\Sigma|^{-\frac{n+n_\mathrm{f}+m+1}{2}} g\{\mathrm{tr}(\Sigma^{-1}[(B-\beta)XX'(B-\beta)' + S + (B_\mathrm{f}-\beta)X_\mathrm{f}X_\mathrm{f}'(B_\mathrm{f}-\beta)' + S_\mathrm{f}])\}\, d\Sigma. \tag{2.18}$$

To evaluate the integral in (2.18), we let, $\Sigma^{-1} = \Lambda$, then

$$d\Sigma = |\Lambda|^{-(m+1)}\, d\Lambda.$$

Therefore, we obtain,

$$p(\beta, B_{\mathrm{f}}, S_{\mathrm{f}}|Y, X, X_{\mathrm{f}}) \propto |S_{\mathrm{f}}|^{\frac{n_{\mathrm{f}}-m-p-1}{2}} \int_{\Lambda} |\Lambda|^{\frac{n+n_{\mathrm{f}}-m-1}{2}} g\{\mathrm{tr}(\Lambda[(B-\beta)XX'(B-\beta)' + S + (B_{\mathrm{f}}-\beta)X_{\mathrm{f}}X_{\mathrm{f}}'(B_{\mathrm{f}}-\beta)' + S_{\mathrm{f}}])\}\, d\Lambda.$$

Following Ng [23], we consider G to be a non-singular matrix of order m such that

$$G'G = [(B-\beta)XX'(B-\beta)' + S + (B_{\mathrm{f}}-\beta)X_{\mathrm{f}}X_{\mathrm{f}}'(B_{\mathrm{f}}-\beta)' + S_{\mathrm{f}}].$$

The transformation, $W = G\Lambda G'$ has the Jacobian of the transformation as $|G'G|^{-\frac{m+1}{2}}$. Then integrating the above with respect to W yields the joint density of β, B_{f} and S_{f} as,

$$\begin{aligned} p(\beta, B_{\mathrm{f}}, S_{\mathrm{f}}|Y, X, X_{\mathrm{f}}) &\propto |S_{\mathrm{f}}|^{\frac{n_{\mathrm{f}}-m-p-1}{2}} [(B-\beta)XX'(B-\beta)' + S + (B_{\mathrm{f}}-\beta)X_{\mathrm{f}}X_{\mathrm{f}}'(B_{\mathrm{f}}-\beta)' \\ &\quad + S_{\mathrm{f}}]^{-\frac{n+n_{\mathrm{f}}}{2}} \int_{\Sigma} g\{\mathrm{tr}(W)\}|W|^{\frac{n+n_{\mathrm{f}}-m-1}{2}}\, dW \\ &\propto |S_{\mathrm{f}}|^{\frac{n_{\mathrm{f}}-m-p-1}{2}} [(B-\beta)XX'(B-\beta)' \\ &\quad + S + (B_{\mathrm{f}}-\beta)X_{\mathrm{f}}X_{\mathrm{f}}'(B_{\mathrm{f}}-\beta)' + S_{\mathrm{f}}]^{-\frac{n+n_{\mathrm{f}}}{2}}. \end{aligned} \tag{2.19}$$

The density function in (2.19) can further be expressed as

$$\begin{aligned} p(\beta, B_{\mathrm{f}}, S_{\mathrm{f}}|Y, X, X_{\mathrm{f}}) &\propto |S_{\mathrm{f}}|^{\frac{n_{\mathrm{f}}-m-p-1}{2}} [(\beta - FA^{-1})A(\beta - FA^{-1})' + S \\ &\quad + (B_{\mathrm{f}} - B)M^{-1}(B_{\mathrm{f}} - B)' + S_{\mathrm{f}}]^{-\frac{n+n_{\mathrm{f}}}{2}}, \end{aligned} \tag{2.20}$$

where $F = BXX' + B_{\mathrm{f}}X_{\mathrm{f}}X_{\mathrm{f}}'$, $A = XX' + X_{\mathrm{f}}X_{\mathrm{f}}'$ and $M = [XX']^{-1} + [X_{\mathrm{f}}X_{\mathrm{f}}']^{-1}$.

The joint density function of B_{f} and S_{f} is obtained by integrating out β using matrix-t argument (see [25, p. 139]) from (2.20) as

$$\begin{aligned} p(B_{\mathrm{f}}, S_{\mathrm{f}}|Y, X, X_{\mathrm{f}}) &\propto \int_{\beta} p(\beta, B_{\mathrm{f}}, S_{\mathrm{f}}|Y, X, X_{\mathrm{f}})\, d\beta \\ &\propto |S_{\mathrm{f}}|^{\frac{n_{\mathrm{f}}-m-p-1}{2}} \int_{\beta} [(\beta - FA^{-1})A(\beta - FA^{-1})' \\ &\quad + S + (B_{\mathrm{f}} - B)M^{-1}(B_{\mathrm{f}} - B)' + S_{\mathrm{f}}]^{-\frac{n+n_{\mathrm{f}}}{2}}\, d\beta \\ &\propto |S_{\mathrm{f}}|^{\frac{n_{\mathrm{f}}-m-p-1}{2}} [S + (B_{\mathrm{f}}-B)M^{-1}(B_{\mathrm{f}}-B)' + S_{\mathrm{f}}]^{-\frac{n+n_{\mathrm{f}}-p}{2}}. \end{aligned} \tag{2.21}$$

Finally, the predictive distribution of the future regression matrix B_{f} is obtained as

$$p(B_{\mathrm{f}}|Y, X, X_{\mathrm{f}}) = \frac{\Gamma_m(n/2)|H|^{\frac{m}{2}}}{\pi^{mp/2}\Gamma_m((n-p)/2)} |S|^{-\frac{p}{2}} |I_m + S^{-1}(B_{\mathrm{f}}-B)M^{-1}(B_{\mathrm{f}}-B)'|^{-\frac{n}{2}}, \tag{2.22}$$

which is a matrix-t density. Thus the predictive distribution of the future regression matrix for given data is an $m \times p$ dimensional matrix-t distribution with $(n - p - m + 1)$ degrees of freedom. That is

$$B_{\mathrm{f}} \sim t_{m \times n_{\mathrm{f}}}(B, M, S_Y, n - p - m + 1).$$

This result coincides with that of Haq [13], where he considered the matrix normal error and that of Khan [14], where he considered the matrix-t error. Thus the predictive distribution of future regression matrix is also unaffected by departures from normality or dependent but uncorrelated assumptions to elliptically contoured distribution.

Note that, if we integrate out B_{f} from (2.21), we obtain the predictive distribution of future sum of squares and product (SSP), which is a generalized Beta density with degrees of freedom $(n_{\mathrm{f}} - p)$ and $(n - p)$. Then the result will coincide with that of Khan [14], where he considered the matrix-t distribution for the multivariate linear model. Thus the predictive distribution of future SSP matrix is also unaffected by departures from matrix-t assumptions to elliptically contoured distribution.

3. Application of matrix-t distribution

In this section, we will discuss about the application of matrix-t distribution, which will also be applicable to the predictive distribution of future response as well as future regression matrices.

3.1. Application in spatial prediction

Risk assessment of air pollution often require to estimate concentration levels at locations where monitoring data are not available, using the data observed at other monitoring sites and possibly at different time periods. This is called the spatial interpolation problem. In this area, the successful application of matrix-t distribution has been found in Sun et al. [26], where the authors considered the method developed by Le et al. [20] and obtained the predictive distribution for the unobserved pollution concentrations (nitrogen dioxide, sulphur dioxide, ozone and sulfate ion) in a spatial filed in the province of Ontario, Canada. They concluded that matrix-t distribution performed well in predicting both concentration level and temporal pattern since the pollutants tend to have high spatial correlation. Their empirical study indicated that the heavier tailed predictive matrix-t distribution is required for the pollutant sulphur dioxide. Le et al. [21] developed a Bayesian approach for spatial and temporal interpolation for the stair case structured monitoring situation for a univariate case or single pollutant. They assumed a Gaussian Generalized Inverted Wishart (GIW) model developed by Brown et al. [3] and obtained the predictive distribution as matrix-t. This GIW prior allows different degrees of freedom to be fitted for individual steps by taking into account the available information from sites at the different steps in the staircase. The theory of Le et al. [21] has been demonstrated by application of the spatial prediction to the ambient ozone field for the Southwestern region of British Columbia. Kibria et al. [17] extended the univariate theory of Le et al. [21] to the multivariate case and thereby obtained an empirical hierarchical Bayesian method for temporal and spatial interpolation using all available “staircase” data. They assumed the responses follow a Gaussian distribution and the corresponding covariance follows a GIW prior distribution. The predictive distribution was obtained as a matrix-t with appropriate parameters. Since the predictive distribution has exact matrix-t, Kibria et al. [17] developed a method of moments approach as an alternative to the much more computationally intensive EM algorithm. Their methodology has been demonstrated by mapping $PM_{2.5}$ field for Philadelphia during the period of May 1992–September 1993. The application of matrix-t distribution has also appeared in the work of Zidek et al. [30, references therein], by predicting average hourly concentration of ambient PM_{10} in Vancouver, British Columbia. Their multivariate approach provide predictions for any given hour to borrow strength through its correlation with adjoining hours. Therefore, matrix-t distribution has significant contributions to interpolate the air pollutants and thereby in the field of environmental risk analysis.

3.2. β-expectation tolerance region for Y_{f}

Following Kibria and Haq [16], we will construct β-expectation tolerance regions for future response matrix Y_{f} as well as future response vector y_{f}. Suppose, $U_1 = ZZ'$ with $Z = T(Y_{\mathrm{f}} - B_Y X_{\mathrm{f}})K$, where T is such that $T'T = S_Y^{-1}$ and K is such that $KK' = I_{n_{\mathrm{f}}} - X'_{\mathrm{f}} A^{-1} X_{\mathrm{f}}$. Then $U = (I + U_1)^{-1} U_1$ has a generalized Beta distribution with n_{f} and $n - m$ degrees of freedom. We note that

$$R_Y = \{U | U < U_{1-\beta}\}, \tag{3.23}$$

is a β-expectation tolerance region for the central $100\beta\%$ of the matrix-t distribution being sampled, where $U_{1-\beta}$ is the point exceeded with probability $1 - \beta$ when using Beta distribution with n_{f} and $n - m$ degrees of freedom. That is R_Y is a tolerance region if $U_{1-\beta}$ is such that

$$\frac{1}{B_p(n_{\mathrm{f}}/2, (n-m)/2)} \int_0^{U_{1-\beta}} |U|^{\frac{n_{\mathrm{f}}-p-1}{2}} |I - U|^{\frac{n-m-p-1}{2}} dU = \beta.$$

Note that for $m = 1$, we obtain the predictive distribution of regression model

$$y_{\mathrm{f}} = \beta X_{\mathrm{f}} + \sigma e_{\mathrm{f}},$$

which is a multivariate Student t distribution with $n - p$ degrees of freedom, location parameter vector $b = yX'(XX')^{-1}$ and scale parameter matrix $s^2 H$, where s^2 is the residual sum of squares of the following regression model:

$$y = \beta X + \sigma e.$$

Then following Kibria and Haq [15], one can construct the β expectation tolerance region for the future response vector y_{f}. Furthermore, from a multivariate Student t distribution, the univariate predictive distribution for a particular future response can be obtained. Then using t distribution, a $(1 - \alpha) \times 100\%$ confidence band can be constructed for any particular response.

4. Concluding remarks

We derived the predictive distribution of future response, regression and SSP matrices under the assumptions of multivariate elliptically contoured distributions that cover various well-known and practically applicable distributions. It has been shown that the predictive distributions under elliptically contoured distributions are identical to those obtained under independent normal errors or matrix-t errors. This gives inference robustness with respect to departures from matrix normal or matrix-t to elliptically contoured distributions. This result also coincides with that of Ng [23], where the author obtained the predictive distribution of future responses for elliptically contoured distribution under both classical and improper Bayesian approaches. Thus, it can be concluded that the predictive distributions for multivariate linear model with elliptically contoured distribution under the improper Bayesian, classical and structural approaches are the same. Some real life applications of matrix-t distribution in the field of spatial prediction have been discussed which indicate that matrix-t distribution has significant contributions in the field of environmental risk analysis.

Acknowledgements

The author is grateful to the Editor and reviewer for their valuable and constructive comments which have improved the presentation and quality of the paper. He is also thankful to Dr. Saralees Nadarajah for carefully reading the revised manuscript.

References

[1] T.W. Anderson, K.T. Fang, Inference in multivariate elliptically contoured distribution based on maximum likelihood, in: K.T. Fang, T.W. Anderson (Eds.), Statistical Inference in Elliptically Contoured and Related Distribution, Allerton Press, New York, 1990, pp. 201–216.
[2] G.E.P. Box, G.C. Tiao, Bayesian Inference in Statistical Analysis, Addision-Wesley, Reading, MA, 1973.
[3] P.J. Brown, N.D. Le, J.V. Zidek, Multivariate spatial interpolation and exposure to air pollutants, Canad. J. Statist. 22 (1994) 489–510.
[4] S. Chib, R.C. Tiwari, S.R. Jammalamadaka, Bayes prediction in regressions with elliptical errors, J. Econometrics 38 (1988) 349–360.
[5] K.-T. Fang, S. Kotz, V.M. Ng, Symmetric Multivariate and Related Distributions, Chapman & Hall, London, 1990.
[6] K.-T. Fang, R. Li, Bayesian statistical inference on elliptical matrix distributions, J. Multivariate Anal. 70 (1999) 66–85.
[7] K.-T. Fang, Y.T. Zhang, Generalized Multivariate Analysis, Springer, New York, 1990.
[8] D.A.S. Fraser, M.S. Haq, Structural probability and prediction for the multivariate model, J. Roy. Statist. Soc. Ser. B 31 (1969) 317–331.
[9] D.A.S. Fraser, K.W. Ng, Multivariate regression analysis with spherical error, in: V.P.R. Krishnaiah (Ed.), Multivariate Analysis, North-Holland Publishing Co., New York, 1980, pp. 369–386.
[10] S. Geisser, Bayesian estimation in multivariate analysis, Ann. Math. Statist. 36 (1965) 150–159.
[11] A.K. Gupta, T. Varga, Elliptical Contoured Models in Statistics, Kluwer Academic Publishers, London, 1993.
[12] I. Guttman, P. Hougaard, Studentization and prediction problems in multivariate multiple regression, Comm. Statist. Theory Methods 14 (1985) 1251–1258.
[13] M.S. Haq, Structural relationships and prediction for the multivariate models, Statist. Hifte 23 (1982) 218–227.
[14] S. Khan, Distribution of sum of squares and products matrices for the generalized multilinear matrix t model, J. Multivariate Anal. 83 (2002) 124–140.
[15] B.M.G. Kibria, M.S. Haq, Predictive inference for the elliptical linear model, J. Multivariate Anal. 68 (1999) 235–249.
[16] B.M.G. Kibria, M.S. Haq, The multivariate linear model with matrix-t error variables, J. Appl. Statist. Sci. 9 (2000) 266–277.

[17] B.M.G. Kibria, L. Sun, J.V. Zidek, N. Le, Bayesian spatial prediction of random space-time fields with application to mapping $PM_{2.5}$ exposure, J. Amer. Statist. Assoc. 97 (2002) 112–124.
[18] H.-M. Kim, B.K. Mallik, A note on Bayesian spatial prediction using the elliptical distribution, Statist. Probab. Lett. 64 (2003) 271–276.
[19] T. Kubokawa, M.S. Srivastava, Robust improvement in estimation of a mean matrix in an elliptically contoured distribution, J. Multivariate Anal. 76 (2001) 138–152.
[20] N.D. Le, W. Sun, J.V. Zidek, Bayesian multivariate spatial interpolation with data missing by design, J. Roy. Statist. Soc. Ser. B 59 (1997) 501–510.
[21] N.D. Le, L. Sun, J.V. Zidek, Spatial prediction and temporal backcasting for environmental fields having monotone data patterns, Canad. J. Statist. 29 (2001) 529–554.
[22] R.H. Loschi, P.L. Iglesian, R.B. Arellano-Valle, Predictivistic characterization of multivariate student t model, J. Multivariate Anal. 85 (2003) 10–23.
[23] V.M. Ng, A note on predictive inference for multivariate elliptically contoured distributions, Comm. Statist. Theory Methods 29 (2000) 477–483.
[24] V.M. Ng, Robust Bayesian inference for seemingly unrelated regression with elliptical errors, J. Multivariate Anal. 83 (2002) 409–414.
[25] J. Press, Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, Robert E. Krieger Publishing Company, Florida, USA, 1982.
[26] W. Sun, N.D. Le, J.V. Zidek, R. Burnett, Assessment of a Bayesian multivariate interpolation approach for health impact studies, Environmetrics 9 (1998) 565–586.
[27] B.C. Sutradhar, M.M. Ali, A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate t model, J. Multivariate Anal. 29 (1989) 155–162.
[28] A. Zellner, Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error terms, J. Amer. Statist. Assoc. 71 (1976) 400–405.
[29] A. Zellner, V.K. Chetty, Prediction and decision problems in regression models from the Bayesian point of view, J. Amer. Statist. Assoc. 60 (1965) 608–616.
[30] J.V. Zidek, L. Sun, N. Le, H. Ozkaynak, Contending with space-time interaction in the spatial prediction of pollution: Vancouver's hourly ambient PM_{10} field, Environmetrics 13 (2002) 595–613.