Journal of Statistical Research *ISSN 0256 - 422 X*
2006, Vol. 40, No. 1, pp. 35-43
Bangladesh

THE PREDICTIVE DISTRIBUTION FOR THE HETEROSCEDASTIC MULTIVARIATE LINEAR MODELS WITH ELLIPTICALLY CONTOURED ERROR DISTRIBUTIONS

B. M. Golam Kibria

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

Email: kibriag@fiu.edu

Saralees Nadarajah

Department of Statistics, University of Nebraska
Lincoln, Nebraska 68583, USA

Email: snadaraj@unlserve.unl.edu

SUMMARY

This paper considers the heteroscedastic multivariate linear model with errors following elliptically contoured distributions. The marginal likelihood function of the unknown covariance parameters and the predictive distribution of future responses have been derived. The predictive distribution obtained is a product of m multivariate Student's t distributions. It is interesting to note that when the models are assumed to have elliptically contoured distributions the marginal likelihood function of the parameters as well as the predictive distribution are identical to those obtained under independently distributed normal errors or dependent but uncorrelated Student's t errors. Therefore, the distribution of future responses is unaffected by a change in the error distribution from the multivariate normal and multivariate t distributions to elliptically contoured distributions. This gives inference robustness with respect to departure from the reference case of independent sampling from the multivariate normal or dependent but uncorrelated sampling from multivariate t distributions to elliptically contoured distributions.

Keywords and phrases: Elliptically Contoured Distributions; Marginal likelihood; Multivariate model; Predictive distribution; Robustness

AMS 2000 Classification: Primary 62J05; Secondary 62H99.

© Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.

1 Introduction

The predictive inference for the linear models has been considered by various researchers: Goldberger (1962) and Hahn (1972) used the classical approach; Tiao and Zellner (1964), Geisser (1965), Zellner and Chetty (1965) and Kibria et al. (2002) used the Bayesian approach, Fraser and Haq (1969) used the structural approach, while Haq and Kibria (1997), Kibria and Haq (1998, 1999) and very recently Kibria (2006) used the structural relation of the model for the derivation of the predictive distribution. The error terms in linear models are assumed to be normally and independently distributed in most applied as well as theoretical research work. However, such assumptions may not be appropriate in many practical situations (for examples, see Zellner (1976), Gnanadesikan (1977), Kibria (1996) and Kibria and Haq (1998)). It happens particularly if the error distribution has heavier tails. One can tackle such situations by using the well known t distribution as it has heavier tails than the normal distribution, especially for smaller degrees of freedom (e.g. Fama (1965) and Blatberg and Gonedes (1974)).

The literature on the predictive inference for the heteroscedastic multivariate linear model is limited. Kibria (1999) considered the predictive inference for the heteroscedastic multivariate linear model under the normality assumption and obtained the predictive distribution as a product of m multivariate Student's t distributions. Kibria (2002) considered the predictive inference for future responses under the multivariate t errors and obtained the predictive distribution as a product of m multivariate Student's t distributions. Therefore, the distribution of future responses for a heteroscedastic model is unaffected by a change in the error distribution from the multivariate normal to the multivariate t distribution. The invariance of the predictive distribution for the future responses suggests that the predictive distribution would be invariant to a wide class of error distributions. In this paper a very general assumption is employed, namely that error terms have a multivariate elliptically contoured distribution. The class of elliptically contoured distributions includes various distributions: the multivariate normal, matric t, multivariate Student's t, multivariate Kotz type and multivariate Cauchy. The class of mixtures of normal distributions is a subclass of elliptical distributions as well as the class of spherically symmetric distributions (Fang et al, 1990).

Elliptically contoured distributions for traditional multivariate regression models have been discussed extensively by Anderson and Fang (1990) and Kubokawa and Srivastava (2003) among others. These distributions have also been considered by Chib et al. (1988), Kibria and Haq (1999) and Kibria (2003) in the context of predictive inference for the general linear or multivariate linear model but not for the heteroscedastic multivariate model. In this paper, we show that when the errors of the model (2.1) are assumed to have an elliptically contoured distribution, the prediction distribution of future responses is a product of m multivariate Student's t distributions. Therefore, the assumption of normality as well as multivariate t is robust to the deviation in the direction of elliptical distributions as far as predictive inference is concerned.

The organization of this paper is as follows: The multivariate linear model and the

marginal likelihood function of the unknown covariance matrix are discussed in section 2. The predictive distribution for the future responses are derived in section 3. Some concluding remarks are given in section 4.

2 Multivariate Linear Model and Marginal Likelihood Function

Consider the following multivariate linear model

$$\mathbf{Y} \quad = \quad \mathbf{BX} + \mathbf{\Gamma E}, \tag{2.1}$$

where $\mathbf{Y}$ and $\mathbf{X}$ are the $m \times n$ and $p \times n$ $(n \geq p)$ response and regressor matrices, respectively, $\mathbf{B}$ is an $m \times p$ matrix of regression parameters, $\mathbf{E}$ is an $m \times n$ errors matrix, and $\mathbf{\Gamma} = \text{diag}\big(\sigma_1, \sigma_2 \ldots, \sigma_{\text{m}}\big)$ is an $m \times m$ scale parameter matrix. Here each of the m rows of $\mathbf{Y}$ may be viewed as a realization of the linear model. When the diagonal elements of $\mathbf{\Gamma}$ are not equal, the model (2.1) is known as heteroscedastic multivariate linear model. Assume that the covariance matrix of each row of $\mathbf{E}$ is an $n \times n$ matrix $\mathbf{\Omega_\lambda}$, where $\boldsymbol{\lambda}$ is a function of a set of unknown parameters. Then the covariance matrix of $\mathbf{E}$ is $\mathbf{\Omega_\lambda} \bigotimes \mathbf{I}_m$, where $\bigotimes$ is the kronecker product between two matrices and $\mathbf{I}_m$ is the identity matrix of order m. The covariance matrix of $\mathbf{Y}$ is $\mathbf{\Omega_\lambda} \bigotimes \mathbf{\Sigma}$, where $\mathbf{\Sigma} = \mathbf{\Gamma\Gamma'}$. The application of model (2.1) has been discussed by Parthasaradhi (1972), Delgado (1992) and Kibria (1999); to mention a few.

We assume that $\mathbf{E}$ has an elliptically contoured distribution with the probability density function

$$f(\mathbf{E}|\mathbf{\Omega_\lambda}) \quad \propto \quad \big|\mathbf{\Omega_\lambda}\big|^{-\frac{m}{2}} g\{tr(\mathbf{E\Omega_\lambda^{-1}E'})\}, \tag{2.2}$$

which is of the form given in Anderson and Fang (1990), where $g\{\cdot\}$ is a non-negative function over $m \times m$ positive definite matrices such that $f(\mathbf{E})$ is a density function. Here, $\mathbf{E}'$ denotes the transpose of the matrix $\mathbf{E}$ and $tr(\mathbf{M})$ denotes the trace of the matrix $\mathbf{M}$.

To derive the marginal likelihood function of $\mathbf{\Omega_\lambda}$, we define the following:

$$\begin{aligned}
\mathbf{\hat{B}_E} &= \mathbf{EX'}(\mathbf{XX'})^{-1}, \\
\mathbf{S_E} &= \text{diag}(\mathbf{E} - \mathbf{\hat{B}_E X})(\mathbf{E} - \mathbf{\hat{B}_E X})' \\
&= \text{diag}(\text{s}^2_{\mathbf{e}1}, \text{s}^2_{\mathbf{e}2}, \ldots, \text{s}^2_{\mathbf{e}\text{m}}), \\
\mathbf{Z_E} &= \mathbf{C_E^{-1}}\Big\{\mathbf{E} - \mathbf{\hat{B}_E X}\Big\},
\end{aligned} \tag{2.3}$$

where $\mathbf{C_E} = \text{diag}(\text{s}_{\mathbf{e}1}, \ldots, \text{s}_{\mathbf{e}\text{m}})$. The corresponding expressions for the response matrix $\mathbf{Y}$ will be denoted by $\mathbf{\hat{B}_Y}$, $\mathbf{S_Y}$ and $\mathbf{Z_Y}$, respectively. ¿From (2.1) and (2.3) we obtain

$$\begin{aligned}
\mathbf{\Gamma} &= \mathbf{C_Y C_E^{-1}}, \\
\mathbf{B} &= \mathbf{\hat{B}_Y} - \mathbf{C_Y C_E^{-1}\hat{B}_E}, \\
\mathbf{Z_E} &= \mathbf{Z_Y}.
\end{aligned} \tag{2.4}$$

The i^{th} row vector of the matrix $\mathbf{E}$ can be written as

$$\mathbf{e}_i = \mathbf{b}_{\mathbf{e}i}\mathbf{X} + s_{\mathbf{e}i}\mathbf{z}_{\mathbf{e}i}, \quad i = 1, 2 \ldots, m, \tag{2.5}$$

where $\mathbf{b}_{\mathbf{e}i}$ is the i^{th} row vector of $\hat{\mathbf{B}}_{\mathbf{E}}$, $\mathbf{z}_{\mathbf{e}i}$ is the i^{th} row vector of $\mathbf{Z}_{\mathbf{E}}$ and $s_{\mathbf{e}i}$ is the i^{th} diagonal element of $\mathbf{C}_{\mathbf{E}}$. Then the relationship between the volume elements of $\mathbf{e}_i$ can be expressed in terms of new variables $\mathbf{b}_{\mathbf{e}i}$, $s_{\mathbf{e}i}$ and $\mathbf{z}_{\mathbf{e}i}$ as

$$d\mathbf{e}_i = |\mathbf{X}\mathbf{X}'|^{\frac{1}{2}} d\mathbf{b}_{\mathbf{e}i} s_{\mathbf{e}i}^{n-p-1} ds_{\mathbf{e}i} d\mathbf{z}_{\mathbf{e}i}, \quad i = 1, 2 \ldots, m, \tag{2.6}$$

(see Fraser and Ng (1980, p. 381)). Since the predictive distribution depends on $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$, we will derive the marginal likelihood function of $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$ first.

Using (2.3) and the Jacobian (2.6), we have the joint density function of $\hat{\mathbf{B}}_{\mathbf{E}}$, $\mathbf{C}_{\mathbf{E}}$ and $\mathbf{Z}_{\mathbf{E}}$ for given $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$ as

$$\begin{aligned} p(\hat{\mathbf{B}}_{\mathbf{E}}, \mathbf{C}_{\mathbf{E}}, \mathbf{Z}_{\mathbf{E}} | \boldsymbol{\Omega}_{\boldsymbol{\lambda}}) \quad \propto \quad & |\boldsymbol{\Omega}_{\boldsymbol{\lambda}}|^{-\frac{m}{2}} \prod_{i=1}^{m} s_{\mathbf{e}i}^{n-p-1} g\Big[\Big\{ \sum_{i=1}^{m} (\mathbf{b}_{\mathbf{e}i} + s_{\mathbf{e}i}\mathbf{g}'_{\mathbf{e}i}\mathbf{F}^{-1}) \\ & \times \mathbf{F}(\mathbf{b}_{\mathbf{e}i} + s_{\mathbf{e}i}\mathbf{g}'_{\mathbf{e}i}\mathbf{F}^{-1})' + \sum_{i=1}^{m} s_{\mathbf{e}i}^2 \mathbf{z}_{\mathbf{e}i} \boldsymbol{\Lambda}_{\boldsymbol{\lambda}} \mathbf{z}'_{\mathbf{e}i} \Big\} \Big], \end{aligned} \tag{2.7}$$

where $\mathbf{F} = \mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}\mathbf{X}'$, $\mathbf{g}_{e_i}$ is the i^{th} row vector of the matrix $\mathbf{G} = \mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}\mathbf{Z}'_{\mathbf{E}}$, and

$$\boldsymbol{\Lambda}_{\boldsymbol{\lambda}} = \boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1} - \boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}\mathbf{X}')^{-1}\mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}.$$

Now, we integrate out $\hat{\mathbf{B}}_{\mathbf{E}}$ and $\mathbf{C}_{\mathbf{E}}$ from (2.7), and obtain the marginal probability density of $\mathbf{Z}_{\mathbf{E}}$ for given $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$. However, following equation (2.4) the *pdf* of $\mathbf{Z}_{\mathbf{Y}}$ can easily be obtained from the *pdf* of $\mathbf{Z}_{\mathbf{E}}$. Therefore, the marginal likelihood function of $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$ conditioned on $\mathbf{Y}$ is given by

$$p(\boldsymbol{\Omega}_{\boldsymbol{\lambda}} | \mathbf{Z}_{\mathbf{Y}}) \quad \propto \quad \prod_{i=1}^{m} |\boldsymbol{\Omega}_{\boldsymbol{\lambda}}|^{-\frac{1}{2}} |\mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}\mathbf{X}'|^{-\frac{1}{2}} \left[\mathbf{z}_{\mathbf{y}i} \boldsymbol{\Lambda}_{\boldsymbol{\lambda}} \left(\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1} - \boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1}\mathbf{X}')^{-1}\mathbf{X}\boldsymbol{\Omega}_{\boldsymbol{\lambda}}^{-1} \right) \mathbf{z}'_{\mathbf{y}i} \right]^{-\frac{n-p}{2}}. \tag{2.8}$$

For a given observation matrix $\mathbf{Y}$ and known design matrix $\mathbf{X}$, the maximum likelihood estimate of $\boldsymbol{\lambda}$ and hence $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$ may be obtained by maximizing the likelihood function (2.8). It appears that a closed form estimate for the parameters may not be available from (2.8). However, for a given observation matrix $\mathbf{Y}$, the MLE $\hat{\boldsymbol{\lambda}}$ of $\boldsymbol{\lambda}$ may be obtained numerically. The marginal likelihood function in (2.8) is identical to that obtained under the assumption of independently distributed multivariate normal errors (see Kibria (1997)) and Student's t errors (see Kibria (2002)).

3 Predictive Distribution

Consider a set of n_f future responses from (2.1) corresponding to the design matrix $\mathbf{X}_f$ as

$$\mathbf{Y}_f \quad = \quad \mathbf{B}\mathbf{X}_f + \mathbf{\Gamma}\mathbf{E}_f, \tag{3.1}$$

where $\mathbf{Y}_f$ and $\mathbf{E}_f$ are the $m \times n_f$ matrices of future responses and errors, respectively, and $\mathbf{X}_f$ is an $p \times n_f$ matrix of future regressors. To derive the joint density of $\mathbf{E}$ and $\mathbf{E}_f$, we combine the observed and future error matrices as

$$\mathbf{E}^* \quad = \quad (\mathbf{E}, \mathbf{E}_f)\,,$$

where $\mathbf{E}^*$ is an $m \times (n + n_f)$ matrix with $Cov(\mathbf{E}^*) = \mathbf{\Phi}_{\boldsymbol{\lambda}} \bigotimes \mathbf{I}_m$, where

$$\mathbf{\Phi}_{\boldsymbol{\lambda}} \quad = \quad \begin{pmatrix} \mathbf{\Phi}_{\boldsymbol{\lambda}11} & \mathbf{\Phi}_{\boldsymbol{\lambda}12} \\ \mathbf{\Phi}_{\boldsymbol{\lambda}21} & \mathbf{\Phi}_{\boldsymbol{\lambda}22} \end{pmatrix},$$

where $\mathbf{\Phi}_{\boldsymbol{\lambda}11}$ is an $n \times n$ covariance matrix of $\mathbf{e}_i$, $\mathbf{\Phi}'_{\boldsymbol{\lambda}12} = \mathbf{\Phi}_{\boldsymbol{\lambda}21}$ is an $n \times n_f$ matrix of covariances between the components of $\mathbf{e}_i$ and $\mathbf{e}_{fi}$ and $\mathbf{\Phi}_{\boldsymbol{\lambda}22}$ is an $n_f \times n_f$ covariance matrix of $\mathbf{e}_{fi}$, $i = 1, \ldots, m$. Following Muirhead (1982), the inverse of $\mathbf{\Phi}_{\boldsymbol{\lambda}}$ is obtained as

$$\mathbf{\Phi}_{\boldsymbol{\lambda}}^{-1} \quad = \quad \begin{pmatrix} \mathbf{\Phi}_{\boldsymbol{\lambda}}^{11} & \mathbf{\Phi}_{\boldsymbol{\lambda}}^{12} \\ \mathbf{\Phi}_{\boldsymbol{\lambda}}^{21} & \mathbf{\Phi}_{\boldsymbol{\lambda}}^{22} \end{pmatrix}.$$

Assume that $\mathbf{E}^*$ has a multivariate elliptically contoured distribution with a known covariance matrix $\mathbf{\Phi}_{\boldsymbol{\lambda}}$. Then we have,

$$p(\mathbf{E}, \mathbf{E}_f | \mathbf{\Phi}_{\boldsymbol{\lambda}}) \quad \propto \quad |\mathbf{\Phi}_{\boldsymbol{\lambda}}|^{-\frac{m}{2}} g\Big[tr\Big\{\mathbf{E}\mathbf{\Phi}_{\boldsymbol{\lambda}}^{11}\mathbf{E}' + \mathbf{E}\mathbf{\Phi}_{\boldsymbol{\lambda}}^{12}\mathbf{E}'_f + \mathbf{E}_f\mathbf{\Phi}_{\boldsymbol{\lambda}}^{21}\mathbf{E}' + \mathbf{E}_f\mathbf{\Phi}_{\boldsymbol{\lambda}}^{22}\mathbf{E}'_f\Big\}\Big].$$

Using (2.3) and the Jacobian (2.6), the joint density function of $\hat{\mathbf{B}}_{\mathbf{E}}$, $\mathbf{C}_{\mathbf{E}}$ and $\mathbf{E}_f$ for given $\mathbf{Z}_{\mathbf{Y}}$ and $\mathbf{\Phi}_{\boldsymbol{\lambda}}$ is obtained as

$$\begin{aligned} & p(\hat{\mathbf{B}}_{\mathbf{E}}, \mathbf{C}_{\mathbf{E}}, \mathbf{E}_f | \mathbf{Z}_{\mathbf{Y}}, \mathbf{\Phi}_{\boldsymbol{\lambda}}) \\ \propto \quad & |\mathbf{\Phi}_{\boldsymbol{\lambda}}|^{-\frac{m}{2}} \prod_{i=1}^{m} s_{\mathbf{e}i}^{n-p-1} g\Big[tr\Big\{(\hat{\mathbf{B}}_{\mathbf{E}}\mathbf{X} + \mathbf{C}_{\mathbf{E}}\mathbf{Z}_{\mathbf{E}})\mathbf{\Phi}_{\boldsymbol{\lambda}}^{11}(\hat{\mathbf{B}}_{\mathbf{E}}\mathbf{X} + \mathbf{C}_{\mathbf{E}}\mathbf{Z}_{\mathbf{Y}})' + (\hat{\mathbf{B}}_{\mathbf{E}}\mathbf{X} + \mathbf{C}_{\mathbf{E}}\mathbf{Z}_{\mathbf{Y}})\mathbf{\Phi}_{\boldsymbol{\lambda}}^{12}\mathbf{E}'_f \\ & + \mathbf{E}_f\mathbf{\Phi}_{\boldsymbol{\lambda}}^{21}(\hat{\mathbf{B}}_{\mathbf{E}}\mathbf{X} + \mathbf{C}_{\mathbf{E}}\mathbf{Z}_{\mathbf{Y}})' + \mathbf{E}_f\mathbf{\Phi}_{\boldsymbol{\lambda}}^{22}\mathbf{E}'_f\Big\}\Big]. \end{aligned} \tag{3.2}$$

Consider the following transformation:

$$\begin{cases} \mathbf{D} = \mathbf{C}_{\mathbf{E}}^{-1}(\mathbf{E}_f - \hat{\mathbf{B}}_{\mathbf{E}}\mathbf{X}_f), \\ \mathbf{U} = \hat{\mathbf{B}}_{\mathbf{E}}, \\ \mathbf{V} = \mathbf{C}_{\mathbf{E}}. \end{cases}$$

The Jacobian of the transformations $J\{(\hat{\mathbf{B}}_{\mathbf{E}}, \mathbf{C}_{\mathbf{E}}, \mathbf{E}_f) \rightarrow (\mathbf{U}, \mathbf{V}, \mathbf{D})\}$ is equal to $\prod_{i=1}^{m} v_i^{n_f}$, where v_i is the i^{th} diagonal element of the matrix $\mathbf{V}$. Then the joint density function of $\mathbf{D}$, $\mathbf{V}$ and $\mathbf{U}$ for given $\mathbf{Z_Y}$ and $\mathbf{\Phi_\lambda}$ becomes

$$\begin{aligned} & p(\mathbf{D}, \mathbf{U}, \mathbf{V} | \mathbf{Z_Y}, \mathbf{\Phi_\lambda}) \\ \propto \quad & |\mathbf{\Phi_\lambda}|^{-\frac{m}{2}} |\mathbf{H}|^{-\frac{m}{2}} \prod_{i=1}^{m} v_i^{n+n_f-p-1} \\ & \times g\left\{ \left(\sum_{i=1}^{m} (\mathbf{u}_i + v_i \mathbf{m}_i \mathbf{H}^{-1}) \mathbf{H} (\mathbf{u}_i + v_i \mathbf{m}_i \mathbf{H}^{-1})' + \sum_{i=1}^{m} v_i^2 r_{\mathbf{\lambda}ii} \right) \right\}, \end{aligned} \tag{3.3}$$

where $\mathbf{u}_i$ is the i^{th} row vector of the matrix $\mathbf{U}$, $\mathbf{m}_i$ is the i^{th} row vector of the matrix $\mathbf{M} = \mathbf{P}_1 \mathbf{Z_Y}' + \mathbf{P}_2 \mathbf{D}'$, $\mathbf{P}_1 = \mathbf{X}\mathbf{\Phi}_{\mathbf{\lambda}}^{11} + \mathbf{X}\mathbf{\Phi}_{\mathbf{\lambda}}^{12}$, $\mathbf{P}_2 = \mathbf{X}_f \mathbf{\Phi}_{\mathbf{\lambda}}^{21} + \mathbf{X}_f \mathbf{\Phi}_{\mathbf{\lambda}}^{22}$ and $r_{\mathbf{\lambda}ii}$ is the i^{th} diagonal element of $\mathbf{R_\lambda} = \mathbf{A} - \mathbf{M}'\mathbf{H^{-1}}\mathbf{M}$, where $\mathbf{A} = \mathbf{Z_Y}\mathbf{\Phi}_{\mathbf{\lambda}}^{11}\mathbf{Z_Y}' + \mathbf{Z_Y}\mathbf{\Phi}_{\mathbf{\lambda}}^{12}\mathbf{D}' + \mathbf{D}\mathbf{\Phi}_{\mathbf{\lambda}}^{21}\mathbf{Z_Y}' + \mathbf{D}\mathbf{\Phi}_{\mathbf{\lambda}}^{22}\mathbf{D}'$ and $\mathbf{H} = \mathbf{X}\mathbf{\Phi}_{\mathbf{\lambda}}^{11}\mathbf{X}' + \mathbf{X}\mathbf{\Phi}_{\mathbf{\lambda}}^{12}\mathbf{x}_f' + \mathbf{X}_f\mathbf{\Phi}_{\mathbf{\lambda}}^{21}\mathbf{X}' + \mathbf{X}_f\mathbf{\Phi}_{\mathbf{\lambda}}^{22}\mathbf{X}_f'$. After integrating out both $\mathbf{U}$ and $\mathbf{V}$ from (3.3), the marginal *pdf* of $\mathbf{D}$ for given $\mathbf{Z_Y}$ and $\mathbf{\Phi_\lambda}$ is obtained as

$$p(\mathbf{D}|\mathbf{Z_Y}, \mathbf{\Phi_\lambda}) \quad \propto \quad |\mathbf{\Phi}(\boldsymbol{\lambda})|^{-\frac{m}{2}} |\mathbf{H}|^{-\frac{m}{2}} \prod_{i=1}^{m} r_{\mathbf{\lambda}ii}^{-\frac{n+n_f-p}{2}}.$$

The i^{th} diagonal element of $\mathbf{R_\lambda}$ can be expressed

$$\mathbf{z}_{\mathbf{y}i}\mathbf{Q}^*\mathbf{z}_{\mathbf{y}i}' + (\mathbf{d}_i + \mathbf{z}_{\mathbf{y}i}\mathbf{Q}_3\mathbf{Q}_2{}^{-1})\mathbf{Q}_2(\mathbf{d}_i + \mathbf{z}_{\mathbf{y}i}\mathbf{Q}_3\mathbf{Q}_2{}^{-1})',$$

where $\mathbf{Q}_1 = \mathbf{\Phi}_{\mathbf{\lambda}}^{11} - \mathbf{P}_1'\mathbf{H}^{-1}\mathbf{P}_1$, $\mathbf{Q}_2 = \mathbf{\Phi}_{\mathbf{\lambda}}^{22} - \mathbf{P}_2'\mathbf{H}^{-1}\mathbf{P}_2$, $\mathbf{Q}_3 = \mathbf{\Phi}_{\mathbf{\lambda}}^{12} - \mathbf{P}_1'\mathbf{H}^{-1}\mathbf{P}_2$ and $\mathbf{Q}^* = \mathbf{Q}_1 - \mathbf{Q}_3\mathbf{Q}_2^{-1}\mathbf{Q}_3'$. Finally, using the relationship $s_{\mathbf{e}i}^{-1}(\mathbf{e}_{fi} - \mathbf{b}_{\mathbf{e}i}\mathbf{X}_f) = s_{\mathbf{y}i}^{-1}(\mathbf{y}_{fi} - \mathbf{b}_{\mathbf{y}i}\mathbf{X}_f) = \mathbf{d}_i$ for $i = 1, \ldots, m$, we obtain the *pdf* of $\mathbf{Y}_f$ for given $\mathbf{Y}$ and $\mathbf{\Phi_\lambda}$ as:

$$\begin{aligned} p\left(\mathbf{Y}_f | \mathbf{Y}, \mathbf{\Phi_\lambda}\right) \quad = \quad & \frac{\left(\Gamma(\frac{n-p}{2})\right)^m}{\left(\Gamma\left(\frac{n-p+n_f}{2}\right)\right)^m} \prod_{i=1}^{m} |\mathbf{\Delta}_{\mathbf{\lambda}i}|^{-\frac{1}{2}} \\ & \times \prod_{i=1}^{m} \left[1 + \frac{1}{n-p}\left\{(\mathbf{y}_{fi} - \boldsymbol{\mu}_{\mathbf{\lambda}i})\mathbf{\Delta}_{\mathbf{\lambda}i}(\mathbf{y}_{fi} - \boldsymbol{\mu}_{\mathbf{\lambda}i})'\right\}\right]^{-\frac{n-p+n_f}{2}} \end{aligned} \tag{3.4}$$

where $\boldsymbol{\mu}_{\mathbf{\lambda}i} = \mathbf{b}_{\mathbf{y}i}\mathbf{X}_f - s_{\mathbf{y}i}\mathbf{z}_{\mathbf{y}i}\mathbf{Q}_3\mathbf{Q}_2^{-1}$ and $\mathbf{\Delta}_{\mathbf{\lambda}i} = \left\{\frac{s_{\mathbf{y}i}^2}{n-p}[\mathbf{z}_{\mathbf{y}i}\mathbf{Q}^*\mathbf{z}_{\mathbf{y}i}']\right\}^{-1}\mathbf{Q}_2$. This is the product of m multivariate Student's t distributions. Properties of this distribution can be found in Dickey (1967) and possible applications can be found in Kibria et al. (2002).

It is observed that for known $\boldsymbol{\lambda}$ and hence $\mathbf{\Phi_\lambda}$, each row of $\mathbf{Y}_f$ is distributed as a n_f dimensional multivariate Student's t distribution with $(n-p)$ degrees of freedom. The location parameter vector is $\boldsymbol{\mu}_{\mathbf{\lambda}i}$ and the scale parameter matrix is $\mathbf{\Delta}_{\mathbf{\lambda}i}^{-1}$ for $i = 1, 2, \ldots, m$. The predictive distribution in (3.4) is identical to that obtained under the assumption of independently distributed multivariate normal errors (see Kibria (1999)) and Student's t error (see Kibria (2002)).

4 Concluding Remarks

This paper has derived the predictive distribution for future responses from the heteroscedastic multivariate linear models under the assumption of multivariate elliptically contoured error distributions. The predictive distribution obtained is a product of multivariate Student's t distributions. It is noted that the predictive distributions under elliptically contoured distributions are identical to those obtained under independent normal errors or dependent but uncorrelated Student's t errors. This gives inference robustness with respect to departures from multivariate normal or Student's t to elliptically contoured distributions. This paper has considered the multivariate heteroscedastic linear model, which also cover the linear model for $m = 1$.

References

[1] Anderson, T. W. and Fang, K. T. (1990). Inference in multivariate elliptically contoured distribution based on maximum likelihood. In: *Statistical Inference in Elliptically Contoured and Related Distribution* (editors K. T. Fang and T. W. Anderson), 201–216, Allerton Press, New York.

[2] Blattberg, R. C. and Gonedes, N. J. (1974). A comparison of the stable and Student's t distributions as statistical models for stock prices. *Journal of Business*, **47**, 224–280.

[3] Chib, S., Tiwari, R. C. and Jammalamadaka, S. R. (1988). Bayes prediction in regressions with elliptical errors. *J. Econometrics*, 38, 349-360.

[4] Delgado, M. A. (1992). Semi parametric generalized least squares in the multivariate non-linear regression model. *Econometric Theory*, **8**, 203–222.

[5] Dickey, J. M. (1967). Expansions of t densities and related complete integral. *Annals of Mathematical Statistics*, **38**, 503–510.

[6] Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, **38**, 34–105.

[7] Fang, K-T., Kotz, S. and Ng, V. M. (1990). Symmetric Multivariate and related distributions. *Chapman and Hall*, London.

[8] Fraser, D. A. S. and Haq, M. S. (1969). Structural probability and prediction for the multivariate model. *Journal of the Royal Statistical Society* B, **31**, 317–331.

[9] Fraser, D. A. S. and Ng, K. W.(1980). Multivariate regression analysis with spherical error. In: *In Multivariate analysis V* (editor P. R. Kirshnaiah), pp. 369-386, North-Holland Publishing Co., New York.

[10] Geisser, S. (1965). Bayesian estimation in multivariate analysis. *Annals of Mathematical Statistics*, **36**, 150–159.

[11] Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. Wiley, New York.

[12] Goldberger, A. S. (1962). Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association*, **57**, 369–375.

[13] Hahn, G. J. (1972). Simultaneous prediction intervals for a regression model. *Technometrics*, **14**, 203–214.

[14] Haq, M. S. and Kibria, B. M. G. (1997). Predictive inference for linear and multivariate linear models with MA(1) error processes. *Communications in Statistics* A, **26**, 331–353.

[15] Kibria, B. M. G. (1996). On preliminary test ridge regression estimator for the restricted linear model with non-normal disturbances. *Communications in Statistics—Theory and Methods*, **25**, 2349–2369.

[16] Kibria, B. M. G. (1997). Contribution to predictive inference for the multivariate linear models. *Ph. D. Thesis*, Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Canada.

[17] Kibria, B. M. G. (1999). The prediction distribution for the heteroscedastic multivariate linear models. *Communications in Statistics—Theory and Methods*, **28**, 2477–2486.

[18] Kibria, B. M. G. (2002). Robust inference of the heteroscedastic multivariate linear models with multivariate Student's t errors. *International Mathematics Journal*, **2**, 833–839.

[19] Kibria, B. M. G. (2003). Robust Predictive Inference for the Multivariate Linear Models with Elliptically Contoured Distribution. *Far East Journal of Theoretical Statistics*, 10(1), 11-24.

[20] Kibria, B. M. G. (2006). The Matric t Distribution and its Applications in Predictive Inference. *Journal of Multivariate Analysis*, 97, (3), 785-795.

[21] Kibria, B. M. G. and Haq, M. S. (1998). Marginal likelihood and prediction for the multivariate ARMA (1,1) linear models with multivariate t error distribution. *Journal of Statistical Research*, **32**, 71–80.

[22] Kibria, B. M. G. and Haq, M. S. (1999). Predictive inference for the elliptical linear model. *Journal of Multivariate Analysis*, **68**, 235–249.

[23] Kibria, B. M. G., Sun, L., Zidek, J. V. and Nhu, L. (2002). Bayesian Spatial Prediction of Random Space-Time Fields With Application to Mapping $PM_{2.5}$ Exposure. *Journal of the American Statistical Association*, 97, 112-124.

[24] Kubokawa, T. and Srivastava, M. S. (2003). Estimating the covariance matrix: a new approach. *Journal of Multivariate Analysis*, **86**, 28–47.

[25] Murihead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

[26] Parthasaradhi, M. (1972). Necessary and sufficient for MINQU-estimation of heteroscedastic variances in linear model. **67**, 486–487.

[27] Tiao, G. C. and Zellner, A. (1964). Bayes' theorem and the use of prior knowledge in regression analysis. *Biometrika*, **51**, 219–230.

[28] Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student's t error terms. *Journal of the American Statistical Association*, **71**, 400–405.

[29] Zellner, A. and Chetty, V. K. (1965). Prediction and decision problems in regression models from the Bayesian point of view. *Journal of the American Statistical Association*, **60**, 608–616.