Ng, V.M. (2004) ***Some robust distributions for the structural multilinear model.*** **International Journal of Statistical Sciences, 3s (Special Issue). pp. 61-67.**

International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 3 (Special Issue), 2004, pp 61–67

Some Robust Distributions for the Structural Multilinear Model

Vee Ming Ng
School of Engineering Sciences
Murdoch University, Murdoch 6150, Australia.

[Received April 15, 2004; Accepted August 10, 2004]

Abstract

The structural approach to inference for location parameters and future responses are considered for the multilinear model with elliptical error distribution. We show that the structural and prediction distributions under elliptical errors assumption are the same as those obtained under normally distributed errors. This gives inference robustness with respect to departures from the reference case of normally distributed errors.

In this paper, we show that the assumption of multivariate elliptically contoured error distributions leads to some structural and prediction distributions that are robust with respect to departures of any elliptically contoured density from the reference case of independent sampling from the normal density, that is, these distributions remain unaffected by the change in error distribution from normal to elliptical. Thus, statistical and predictive inference is robust with respect to departures of any elliptically contoured density from the reference case of independent sampling from the normal density. This robustness result holds for a large class of density functions since elliptically contoured distributions include the multivariate normal, matric-t and the multivariate Cauchy.

The structural and prediction distributions are obtained by integrating over the relevent parameter spaces and the prediction distribution so derived extends and generalizes some results of Kibria and Haq (1999) and Khan and Haq (1994). Although the structural distribution of the regression parameters has been considered under parameter and distribution factorization methods in Fraser and Ng (1980), it is derived here for completeness.

In section 2, the structural distribution of regression parameters is derived and in section 3, the predictive distribution is obtained. Some concluding comments are made in section 4.

2 The Structural Model

Consider the multivariate regression model with elliptical errors,

$$Y = \mathcal{B}X + E \tag{1}$$

where Y is a m x n matrix, representing n vectors each of m components; X is a r x n design matrix of known values of rank r, and $\mathcal{B}$ is the m x r matrix of regression parameters, $n > m+r$. The m x n error component E is assumed to have an elliptically contoured distribution with the probability density function

$$f(E|\Sigma) \quad \propto \quad |\Sigma|^{-\frac{n}{2}} h\{\text{tr } (\Sigma^{-1}EE')\} \tag{2}$$

which is of the form given in Fang and Anderson (1990), where $h\{.\}$ is a non-negative function over m x m positive definite matrices such that $f(E|\Sigma)$ is a density function and Σ is a m x m covariance matrix of full rank.

Under the structural approach, the above model is written as the pair:

$$\begin{aligned} &Y = [\mathcal{B}, \Gamma]E \\ &f(E)dm(E) \end{aligned} \tag{3}$$

where

- $[\mathcal{B}, \Gamma]$ is an element of the transformation group, G, that operates on the space of E through the matrix X, producing a observed response Y, in the following manner:

$$Y = [\mathcal{B}, \Gamma]E = \mathcal{B}X + \Gamma E; \tag{4}$$

- the error component has a known probability density function $f(E)$ (which is (2) with Σ equals to the identity matrix) with respect to the invariant probability measure $m(.)$,

- $\Gamma\Gamma^T = \Sigma,\quad \Gamma^T$ is the transpose of Γ and is a mxm matrix.

Following Fraser (1968), a reduced model resulting from the structural model above is:

$$\begin{aligned} t(Y) &= [\mathcal{B}, \Gamma]t(E) \\ k(D)f\{t(E)D\}d\mu(t(E)) \end{aligned} \tag{5}$$

where $t(.)$ is a transformation variable assuming values in the transformation group G such that $t(gE) = gt(E)$ for all g in G and for all E in the sample space, $D = t^{-1}(E)E = t^{-1}(Y)Y$ is a reference point on the orbit of E for a given Y and $k(D)$ is the normalizing constant such that $f\{t(E)D\}$ is a density function with respect to the left invariant measure $\mu(.)$ on G. A suitable transformation variable for the multilinear model is

$$t(E) = [B(E), C(E)] \tag{6}$$

where $B(E) = EX^T(XX^T)^{-1}$ and $S(E) = C(E)C^T(E) = (E-B(E)X)(E-B(E)X)^T$.

A transformation on the space of E of the form $\tilde{E} = gE$ where $g = [B, C]$ yields the Jacobian of transformation $|g|^n = |C|^n$ and the invariant measure $dm(E) = |g|^{-n}dE = |C|^{-n}dE$. A left transformation of the group via $[\tilde{B}, \tilde{C}] = g[B^*, C^*] = [B, C][B^*, C^*] = [B + CB^*, CC^*]$ yields the Jacobian of transformation $|g|^{m+r} = |C|^{m+r}$ and the left invariant measure $d\mu([B, C]) = |C|^{-m-r}dBdC$

The distribution of t(E) describing the error E when expressed in invariant form is

$$f(E)dE = f(E)|t(E)|^n dm(E)$$

The conditional distribution of $t(E)$ given the "observed" orbit $D(E) = D(Y) = D$ is

$$\begin{aligned} &k(D)f(t(E)D)|t(E)|^n d\mu(t(E)) \\ &= k(D)f(t(E)D)|t(E)|^{n-m-r}d(t(E)) \end{aligned} \tag{7}$$

The distribution above describes possible values for $t(E)$ in the equation

$$t(Y) = [\mathcal{B}, \Gamma] t(E)$$

where $t(Y)$ is known and $\mathcal{B}, \Gamma$ are unknown and the distribution reflects the joint distribution of $\mathcal{B}$ and $\Gamma\Gamma^T = \Sigma$ i.e. the structural distribution of $\mathcal{B}, \Sigma$:

$$\begin{aligned} &k(D) f([\mathcal{B}, \Gamma]^{-1} Y) |\Gamma|^{-n-m} |t(Y)|^{n-r} d\mathcal{B} d\Gamma \\ &= k(*) f([\mathcal{B}, \Gamma]^{-1} Y) |\Sigma|^{-\frac{n+m+1}{2}} d\mathcal{B} d\Sigma \end{aligned} \tag{8}$$

$k(*)$ is the normalising constant and is used as a generic symbol for a normalising constant from now on.

For elliptical errors, the joint structural distribution of $\mathcal{B}$ and Σ takes the form

$$k(*) h(\mathrm{tr}(\Sigma^{-1}(B(Y) - \mathcal{B}) X X^T (B(Y) - \mathcal{B})^T + S(Y)) |\Sigma|^{-\frac{n+m+1}{2}} d\mathcal{B} d\Sigma \tag{9}$$

where $B(.)$ and $S(.)$ are defined above.

Let P be a m x m nonsingular matrix such that

$P^T P = (B(Y) - \mathcal{B}) X X^T (B(Y) - \mathcal{B})^T + S(Y)$. The transformation, $W = P \Sigma^{-1} P^T$, has $| P^T P|^{\frac{m+1}{2}}$ as the Jacobian of transformation. Integrating the above with respect to W yields the marginal structural distribution of $\mathcal{B}$:

$$\begin{aligned} &f(\mathcal{B}|Y) \\ &\propto |(B(Y) - \mathcal{B}) X X^T (B(Y) - \mathcal{B})^T + S(Y))|^{-\frac{n}{2}} \int h\{\mathrm{tr}\ W\} |W|^{\frac{n+m+1}{2}} dW^{-1} \\ &\propto |(B(Y) - \mathcal{B}) X X^T (B(Y) - \mathcal{B})^T + S(Y))|^{-\frac{n}{2}} \end{aligned} \tag{10}$$

i.e. the structural distribution of $\mathcal{B}$ is matric-t with $n - m - r + 1$ degrees of freedom: $t_{mr}(B(Y), (X^T X)^{-1}, S(Y), n - r - m + 1)$ using the notation of Box and Tiao (1973, pp. 441-442). The structural distribution of $\mathcal{B}$ under normal errors is also matric-t (Fraser and Haq(1970)) with $n - m - r + 1$ degrees of freedom.

3 Prediction Distribution

Let Y_f be a set of future responses from the multilinear model. Then

$$Y_f = \mathcal{B} X_f + E_f \tag{11}$$

where Y_f is a m x n_f matrix, representing n_f vectors each of m component; X_f is a r x n_f design matrix of known values of rank $n_f \geq r$, and E_f is the m x n_f matrix of future error variable.

The joint density function of (E, E_f) is given by

$$f(E, E_f) \propto h\{\mathrm{tr}\ (E E^T + E_f {E_f}^T)\} \tag{12}$$

From this joint density and the structural relationships, the prediction distribution can be obtained from (see Fraser and Haq (1970))

$$f(Y_f|Y) \propto \int\int |\Sigma|^{-\frac{n_1+n_f+m+1}{2}} \cdot h\{\text{tr } \Sigma^{-1}[(Y - \mathcal{B}X)(Y - \mathcal{B}X)^T + (Y_f - \mathcal{B}X_f)(Y_f - \mathcal{B}X_f)^T]\} d\mathcal{B} d\Sigma \quad (13)$$

We can then rewrite the matrix expression in (13) as

$$\begin{aligned} & (Y - \mathcal{B}X)(Y - \mathcal{B}X)^T + (Y_f - \mathcal{B}X_f)(Y_f - \mathcal{B}X_f)^T \\ = \; & S(Y) + (Y_f - B(Y)X_f)H(Y_f - B(Y)X_f)^T + (\mathcal{B} - B^*)A(\mathcal{B} - B^*)^T \end{aligned}$$

where

$$\begin{aligned} A &= XX^T + X_f X_f{}^T \\ B^* &= (B(Y)XX^T + Y_f X_f{}^T)A^{-1} \\ H &= I_{n_f} - X_f{}^T A^{-1} X_f \\ I_{n_f} & \quad \text{is the } n_f \text{ x } n_f \text{ identity matrix.} \end{aligned}$$

Making the joint transformation

$$\tilde{\mathcal{B}} = \Sigma^{-\frac{1}{2}}(\mathcal{B} - B^*), \quad W = Q\Sigma^{-1}Q^T$$

where $Q^TQ = S(Y) + (Y_f - B(Y)X_f)H(Y_f - B(Y)X_f)^T$ and the Jacobian of transformation $|W|^{-\frac{r}{2}} \mid Q^TQ|^{\frac{m+r+1}{2}}$, (13) becomes

$$\begin{aligned} & f(Y_f|Y) \\ & \quad \propto \int\int |S(Y) + (Y_f - B(Y)X_f)H(Y_f - B(Y)X_f)^T|^{-\frac{n+n_f-r}{2}} |W|^{\frac{n+n_f+m-r+1}{2}} \cdot \\ & \qquad h\{\text{tr } W + \text{tr } \tilde{\mathcal{B}}A\tilde{\mathcal{B}}^T\} d\tilde{\mathcal{B}} dW^{-1} \\ & \quad \propto |S(Y) + (Y_f - B(Y)X_f)H(Y_f - B(Y)X_f)^T|^{-\frac{n+n_f-r}{2}} \end{aligned} \quad (14)$$

Hence Y_f has a matric-t distribution with $n - m - r + 1$ degrees of freedom, i.e. $t_{mn_f}[B(Y)X_f, H^{-1}, S(Y), n-m-r+1]$. The prediction distribution of Y_f is the same as that for normal errors (see Haq and Rinco (1976)). This result extends that in Kibria and Haq (1999) where the case m=1 was considered.

4 Concluding Remarks

When the errors in a regression model are assumed to have an elliptically contoured distribution, the structural distribution of the regression parameters and the predictive distribution are identical to those obtained under independently distributed normal errors. Both distributions involve the matric-t distribution which describes the distribution of "location" variables - the regression parameters and future observations. Hence the structural approach gives uniqueness and robustness to deviation in the direction of elliptical distribution with respect to inference of regression parameters and future observations.

References

[1] Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis.* Addison-Wesley, Reading, Mass.

[2] Fang, K.T. and Anderson, T.W. (1990). "Inference in multivariate ellipti- cally contoured distribution based on maximum likelihood," in *Statist ical Inference in Elliptically Contoured and Related Distribution*, (K.T. Fang and T.W. Anderson Ed.), 201-216. Allerton Press, New York.

[3] Fraser, D.A.S. (1968). *The Structure of Inference.* Wiley, New York.

[4] Fraser, D.A.S and Haq, M. S. (1970) "Inference and Prediction for the Multilinear model", *Statist. Res*, 4 93-109

[5] Fraser, D.A.S. and Ng, K.W. (1980). " Multivariate regression analysis with spherical error," in *Multivariate Analysis-V* (P.R. Krishnaiah Ed.), 369-386. North-Holland, New York.

[6] Geisser, S. (1965). "Bayesian estimation in multivariate analysis," *Ann. Math. Statist.,* **36**, 150-159.

[7] Guttman, I. and Hougaard, P. (1985). "Studentization and prediction problems in multivariate multiple regression," *Commun. Statist. - Theor. Meth.,* **14**, 1251-1258.

[8] Haq, M.S. and Rinco, S. (1976). "β-expectation tolerance regions for a generalized multivariate model with normal error variables," *J. Multivariate Analysis*, **6**, 414-432.

[9] Khan, S. and Haq, M.S. (1994). "β-expectation tolerance region for the multilinear model with matrix-t error distribution," *Commun. Statist. Theor. Meth.,* **23**, 1935-1951.

[10] Kibria, B.M.G. and Haq, M.S. (1999). "Predictive inference for the elliptical linear model," *J. Multivariate Analysis*, **68**, 235-249.

[11] Sutradhar, B.C. and Ali, M. M. (1989). "A generalization of the Wishart distribution for the elliptical model and its moments for the mutivariate t model, " *J. Multivariate Analysis*, **29**, 155-162.

[12] Zellner, A. (1976). "Bayesian and non-Bayesian analysis of the regression model with multivariate student-t error terms," *J. Amer. Statist. Assoc.*, **71**, 400-405.