ESTIMATION OF A SET OF FIXED VARIATES FOR OBSERVED VALUES OF DEPENDENT VARIATES WITH NORMAL MULTIVARIATE REGRESSION MODELS SUBJECTED TO LINEAR RESTRICTIONS

By D. G. Kabe

(Received May 15, 1963)

1. Introduction

Here we consider the problem of estimating a set of fixed variates for observed values of dependent variates with normal multivariate regression models subjected to linear restrictions. Sometimes with regression models we are faced with the problem of estimating the fixed variates by using a set of observed values of dependent variables. This problem does not appear to have drawn much attention. The formulae at present available treat the case of a single fixed variate with univariate normal regression models, see, e.g., Graybill (1961, p. 125). Further the available formulae assume that there are no restrictions on the regression parameters. The formulae also assume that the set of observations on which the estimation of fixed variates is based is independent of the set of observations used for estimating the regression parameters. Our purpose in this paper is to give more general formulae. We treat the case of univariate regression first and then we develop the multivariate case on similar lines. Our aim in doing so is to point out the close analogy that exists between the univariate and the multivariate cases. However, we assume that the set of observations used for estimating the fixed variates is independently distributed of the set of observations used. We may perhaps in a future communication treat the case when these two sets are correlated.

Some results which are found useful in the sequel are stated in the section 2.

2. Some useful results

Sverdrup (1947) gives a lemma which may be stated as follows. Let y be a column vector of N components, D a given $q \times N$ matrix of rank $q(<N)$, then

$$(2.1)\qquad \int_{y'y=u,\ Dy=v} f(y'y,\ Dy)dy$$
$$=\frac{1}{2}C(N-q)\ |\ DD'\ |^{-1/2}f(u,\ v)[u-v'(DD')^{-1}v]^{(N-q-2)/2}\ .$$

Here v is a q component column vector, dy, as usual, denotes the product of the differentials of the elements of y, and $C(N)$ represents the surface area of a unit N dimensional sphere. The integral is considered as a part of the volume integral over the appropriate range. If f is a suitable density function, then the right-hand side of (2.1) obviously represents the joint density of u and the vector variate v. A similar lemma follows when the rank of the matrix D is less than q.

In case the matrix D and the vector v are partitioned into a number of parts, say, two; $D=(D_1'D_2')'$, and $v=(v_1'v_2')'$, then it follows from the lemma that

$$(2.2)\qquad \int_{(y-\mu)'A(y-\mu)'=u,\ (D_1'D_2')'y=(v_1'v_2')'} f((y-\mu)'A(y-\mu))dy$$
$$=\frac{1}{2}C(N-q)\,|A\,|^{-1/2}\,|D_1A^{-1}D_1'\,|^{-1/2}f(u)$$
$$|D_2A^{-1}D_2'-D_2A^{-1}D_1'(D_1A^{-1}D_1')^{-1}D_1A^{-1}D_2'\,|^{-1/2}$$
$$[u-(v_1-\mu_1)'(D_1A^{-1}D_1')^{-1}(v_1-\mu_1)-(v_2-\mu_2$$
$$-D_2A^{-1}D_1'(D_1A^{-1}D_1')^{-1}(v-\mu_1))'(D_2A^{-1}D_2'$$
$$-D_2A^{-1}D_1'(D_1A^{-1}D_1')^{-1}D_1A^{-1}D_2')^{-1}(v_2-\mu_2$$
$$-D_2A^{-1}D_1'(D_1A^{-1}D_1')^{-1}(v_1-\mu_1))]^{(N-q-2)/2}\ .$$

Here A is a $N\times N$ positive definite symmetric matrix of rank N, μ is a column vector of N components, and μ_1 and μ_2 are appropriate parts of the q component column vector $D\mu$. A similar result follows when D and v are partitioned into more than two parts.

Our next result is a generalization of (2.1) given by Kabe (1960). Kabe's result may be stated as follows. Let Y be a $p\times N$ matrix, D a given $q\times N$ matrix of rank $q(<N)$, $N\geqq p+q$, then

$$(2.3)\qquad \int_{YY'=G,\ DY'=V'} f(YY',\ DY')dY$$
$$=2^{-p}\prod_{i=1}^{p}C(N-p-q+i)\,|\,DD'\,|^{-p/2}f(G,\ V')$$
$$|\,G-V(DD')^{-1}V'\,|^{(N-p-q-1)/2}\ .$$

Here G is a $p\times p$ positive definite symmetric matrix of rank p, V is a $p\times q$ matrix, dY, as usual, denotes the product of the differentials of the elements of the matrix Y. The integral is considered as a part of the volume integral over the appropriate range. If f is a suitable density function, then the right-hand side of (2.3) obviously represents the joint density of the matrices G and V.

The lemma and its applications to multivariate normal sampling distribution theory have been recommended for publication in the Annals of Mathematical Statistics and we defer the proof to that publication.

In case D and V are partitioned into a number of parts, say, two: $D=(D_1'D_2')'$, $V=(V_1\ V_2)$, then it follows from (2.3) that

$$\begin{aligned}(2.4)\qquad &\int_{(Y-\mu)A(Y-\mu)'=G,\ (D_1'D_2')'Y'=(V_1'V_2')'} f((Y-\mu)A(Y-\mu)')dY \\ &=2^{-p}\prod_{i=1}^{p}C(N-p-q+i)\,|A|^{-p/2}\,|D_1A^{-1}D_1'|^{-p/2} \\ &|D_2A^{-1}D_2'-D_2A^{-1}D_1'(D_1A^{-1}D_1')^{-1}D_1A^{-1}D_2'|^{-p/2}f(G) \\ &|G-(V_1-\mu_1)(D_1A^{-1}D_1')^{-1}(V_1-\mu_1)'-(V_2-\mu_2 \\ &-(V_1-\mu_1)(D_1A^{-1}D_1')^{-1}D_1A^{-1}D_2')(D_2A^{-1}D_2' \\ &-D_2A^{-1}D_1'(D_1A^{-1}D_1')^{-1}D_1A^{-1}D_2')^{-1}(V_2-\mu_2 \\ &-(V_1-\mu_1)(D_1A^{-1}D_1')^{-1}D_1A^{-1}D_2')'|^{(N-p-q-1)/2}\,.\end{aligned}$$

Here A is a $N\times N$ positive definite symmetric matrix of rank N, μ is a $p\times N$ matrix of constant terms, and μ_1 and μ_2 are appropriate parts of the $p\times q$ matrix $\mu D'$. A similar result obviously follows when D and V are partitioned into more than two parts.

Now we proceed with the estimation problem of fixed variates. We assume that all the integrals occurring in this paper are evaluated over appropriate ranges of the variables of integration.

3. Estimation of fixed variates: Univariate case

We consider a univariate regression model

$$(3.1)\qquad y=X\beta+e\,,$$

where y is a N component column vector, X is a $N\times$q matrix of fixed variates, β is the q component vector of population regression coefficients, and e is the error vector of N components. The vector e has a N variate normal distribution with zero mean vector and covariance matrix $I\sigma^2$, where σ^2 is the unknown variance of each component of e. There is no loss of generality in assuming that the covariance of e to be $I\sigma^2$. Further we suppose that the model is already reparametrized so that we assume the rank of the matrix X to be $q(<N)$. Let now β be subjected to g estimable linear restrictions

$$F\beta=w\,, \tag{3.2}$$

where F is a $g\times q$ matrix of rank $g(<q)$. The least squares (and in our context the maximum likelihood) estimate $\hat{\beta}$ of β is given by the expression

$$\hat{\beta}=(X'X)^{-1}X'y\,. \tag{3.3}$$

The estimated restrictions (3.2) are

$$F\hat{\beta}=\hat{w}\,. \tag{3.4}$$

Further, the vector $HX_0\beta$ is estimated by $HX_0\hat{\beta}$, where the $h\times q$ matrix HX_0 is a known matrix of rank $h(<q)$. We use HX_0 conveniently to denote a different set of values of the matrix X or a part of it. Our problem here now is to find the conditional, given $F\hat{\beta}$, density of the vector $HX_0\hat{\beta}$. It is known that the densities of the variates $\hat{\sigma}^2$, and $\hat{\beta}$, where

$$\hat{\sigma}^2=(y'y-\hat{\beta}'(X'X)\hat{\beta})/(N-q)\,, \tag{3.5}$$

and

$$\hat{\beta}=(X'X)^{-1}X'y\,, \tag{3.6}$$

are independent. The density of $\hat{\sigma}^2$ is

$$g(\hat{\sigma}^2)=2^{-(N-q)/2}(N-q)^{(N-q)/2}\sigma^{-(N-q)}\left(\Gamma\left(\frac{1}{2}\left[N-q\right]\right)\right)^{-1} \exp\left\{-\frac{1}{2\sigma^2}(N-q)\hat{\sigma}^2\right\}. \tag{3.7}$$

The density of $\hat{\beta}$ is given by the expression

$$g(\hat{\beta})=(2\pi)^{-q/2}\sigma^{-q}\,|X'X|^{1/2}\exp\left\{\frac{-1}{2\sigma^2}(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)\right\}. \tag{3.8}$$

Now the joint density of the variates $HX_0\hat{\beta}=\hat{z}$, $\hat{w}$, and $(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)=u_1$, is obviously given by the equation

$$g(u_1,\,\hat{z},\,\hat{w})=\int_R g(\hat{\beta})d\hat{\beta}\,, \tag{3.9}$$

where the region R of integration is determined by the conditions

$$(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)=u_1\,,\quad \begin{pmatrix}F\\ HX_0\end{pmatrix}\hat{\beta}=\begin{pmatrix}\hat{w}\\ \hat{z}\end{pmatrix}. \tag{3.10}$$

The integral is evaluated by using (2.2); and we find that

$$
\begin{aligned}
(3.11)\quad & g(u_1, \hat{z}, \hat{w}) \\
& = (2\pi)^{-q/2}\sigma^{-q}\frac{1}{2}C(q-g-h)\,|F(X'X)^{-1}F'|^{-1/2}\exp\left\{-\frac{1}{2\sigma^2}u_1\right\} \\
& |HX_0(X'X)^{-1}X_0'H' - HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}F(X'X)^{-1}X_0'H'|^{-1/2} \\
& [u_1 - (\hat{w}-w)'(F(X'X)^{-1}F')^{-1}(\hat{w}-w) \\
& -(\hat{z} - HX_0\beta - HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}(\hat{w}-w))' \\
& (HX_0(X'X)^{-1}X_0'H' - HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}F(X'X)^{-1}X_0'H')^{-1} \\
& (\hat{z} - HX_0\beta - HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}(\hat{w}-w))]^{(q-g-h-2)/2}\,.
\end{aligned}
$$

In (3.11) we have assumed that $q \geqq g+h$. It follows that the conditional density of $\hat{z}$, given $\hat{w}$, is h variate normal with mean vector

$$
(3.12)\qquad HX_0\beta + HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}(\hat{w}-w)\,,
$$

and the covariance matrix Σ_1, where

$$
\begin{aligned}
(3.13)\qquad \Sigma_1 = {} & (HX_0(X'X)^{-1}X_0'H' \\
& - HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}F(X'X)^{-1}X_0'H')\sigma^2\,.
\end{aligned}
$$

Now we consider the problem of estimation of fixed variates. The observed vector corresponding to the values $HX_0\beta$ may be denoted by $Hy = v$. The density of v is obviously given by the expression

$$
\begin{aligned}
(3.14)\qquad & g(v) = (2\pi)^{-h/2}\sigma^{-h}\,|HH'|^{-1/2} \\
& \exp\left\{-\frac{1}{2\sigma^2}(v - HX_0\beta)'(HH')^{-1}(v - HX_0\beta)\right\}\,.
\end{aligned}
$$

It is evident that the density of the vector ζ, where

$$
(3.15)\qquad \zeta = \hat{z} - v\,,
$$

is h variate normal with mean vector μ_1, where

$$
(3.16)\qquad \mu_1 = HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}(\hat{w}-w)\,,
$$

and covariance matrix Σ_2, where

$$
\begin{aligned}
(3.17)\qquad \Sigma_2 = {} & (HH' + HX_0(X'X)^{-1}X_0H' \\
& - HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}F(X'X)^{-1}X_0'H')\sigma^2\,.
\end{aligned}
$$

Obviously the above density of ζ is conditional for given $\hat{w}$. It is easily observed that the statistic

$$F_1 = \frac{(\zeta-\mu_2)'\Sigma_2^{-1}(\zeta-\mu_2)/h}{\hat{\sigma}^2/\sigma^2}, \tag{3.18}$$

has a noncentral F distribution with h and $N-q$ degrees of freedom. Here μ_2 is given by

$$\mu_2 = HX_0(X'X)^{-1}F'(F(X'X)^{-1}F')^{-1}\hat{w}. \tag{3.19}$$

The above F distribution may be used to set confidence regions for the vector HX_0.

4. Estimation of fixed variates: Multivariate case

We consider a multivariate regression model

$$Y = BX + E \tag{4.1}$$

where Y is a $p\times N$ matrix, X is a $q\times N$ matrix of fixed variates, B is the $p\times q$ matrix of population regression coefficients, and E is the $p\times N$ matrix of error terms. The pN component vector E has a pN variate normal distribution with mean vector zero, and covariance matrix $\Sigma\otimes I$. We assume our model is already reparametrized and take the rank of the matrix X to be $q(<N)$. Let now B be subjected to g linear estimable restrictions

$$FB' = W', \tag{4.2}$$

where F is a $g\times q$ matrix of rank $g(<q)$. The least squares (and in our context the maximum likelihood) estimate $\hat{B}$ of B is

$$\hat{B}' = (XX')^{-1}XY'. \tag{4.3}$$

The estimated equation (4.2) is

$$F\hat{B}' = \hat{W}'. \tag{4.4}$$

Further, the matrix $HX_0'B'$ is estimated by $HX_0'\hat{B}'$. Here we assume that the $h\times q$ known matrix HX_0' is of rank $h(<q)$. HX_0' conveniently denotes a set of different values of X' or a part of it. Our problem now is to find the conditional, given $\hat{W}$, density of the ph component vector $\hat{B}X_0H'$. It is known that the densities of $\hat{B}$ and $\hat{\Sigma}$, where

$$\hat{\Sigma} = (YY' - \hat{B}(XX')\hat{B}'), \tag{4.5}$$

are independent. $\hat{\Sigma}$ has a central Wishart density with $N-q$ degrees of freedom and population covariance matrix Σ. The density of $\hat{B}$ is

$$g(\hat{B})=|XX'|^{p/2}|\Sigma|^{-q/2}(2\pi)^{-pq/2} \exp\left\{-\frac{1}{2}tr\,\Sigma^{-1}[(\hat{B}-B)(XX')(\hat{B}-B)']\right\}. \tag{4.6}$$

It follows that the joint density of the matrix variates $\hat{W}$, $\hat{Z}$, and $G_1=(\hat{B}-B)(XX')(\hat{B}-B)$, is given by the expression

$$g(\hat{W},\hat{Z},G_1)=\int_R g(\hat{B})d\hat{B}, \tag{4.7}$$

where the region R of integration is determined by the conditions

$$(\hat{B}-B)XX'(\hat{B}-B)'=G_1,\ \begin{pmatrix} F \\ HX_0' \end{pmatrix}\hat{B}'=\begin{pmatrix} \hat{W}' \\ \hat{Z}' \end{pmatrix}. \tag{4.8}$$

The integral is evaluated by using (2.4); and we have that

$$\begin{aligned} g(G_1,\hat{W},\hat{Z}) &= (2\pi)^{-pq/2}|\Sigma|^{-q/2}2^{-p}\prod_{i=1}^{p}C(q-p-g-h+i) \\ &|F(XX')^{-1}F'|^{-p/2}\exp\left\{-\frac{1}{2}tr\,\Sigma^{-1}G_1\right\} \\ &|HX_0'(XX')^{-1}X_0H'-HX_0(XX')^{-1}F'(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H'|^{-p/2} \\ &|G_1-(\hat{W}-W)(F(XX')^{-1}F')^{-1}(\hat{W}-W)' \\ &-(\hat{Z}-BX_0H'-(\hat{W}-W)(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0X') \\ &(HX_0'(XX')^{-1}X_0H' \\ &-HX_0'(XX')^{-1}F'(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H')^{-1}(\hat{Z}-BX_0H' \\ &-(\hat{W}-W)(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H')'|^{(q-p-g-h-1)/2}. \end{aligned} \tag{4.9}$$

In (4.9) we assume that $q\geqq p+g+h$. It follows that the ph variate vector $\hat{Z}$ has a ph variate conditional, given $\hat{W}$, normal distribution with ph component mean vector

$$BX_0H'+(\hat{W}-W)(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H', \tag{4.10}$$

and covariance matrix

$$\Sigma\otimes(HX_0'(XX')^{-1}X_0H' -HX_0'(XX')^{-1}F'(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H'). \tag{4.11}$$

We now proceed with the problem of estimation of the fixed variates. Let the set of observed values corresponding to BX_0H' be denoted by $HY'=V'$, say. Then obviously the density of V' is

$$(4.12)\qquad g(V)=(2\pi)^{-ph/2}\,|\Sigma|^{-h/2}\,|HH'|^{-p/2}\exp\left\{-\frac{1}{2}tr\,\Sigma^{-1}(V-BX_0H')(HH')^{-1}(V-BX_0H')'\right\}.$$

It follows that the density of the ph component vector ζ, where

$$(4.13)\qquad \zeta=\hat{Z}-V,$$

is ph variate normal with mean vector

$$(4.14)\qquad \mu_3=(\hat{W}-W)(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H',$$

and covariance matrix

$$(4.15)\qquad \Sigma\otimes Q,$$

where

$$(4.16)\qquad Q=(HH'+HX_0'(XX')^{-1}X_0H' -HX_0(XX')^{-1}F'(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H').$$

The above density of ζ is conditional for given $\hat{W}$. It follows that the the $p\times p$ positive symmetric definite matrix L, defined by the equations

$$(4.17)\qquad \hat{\Sigma}=CLC',\ (\zeta-\mu_4)Q^{-1}(\zeta-\mu_4)'+\hat{\Sigma}=CC',$$

where C is a lower triangular matrix of order p, has a noncentral multivariate beta distribution as defined by Kshirsagar (1961). Here the ph component vector μ_4 is given by the equation

$$(4.18)\qquad \mu_4=\hat{W}(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H'.$$

The practical applications of the L distribution appear to be limited. However, when $h=1$ the ph component vector ζ reduces to a single p component vector. We denote this p component vector also by ζ. It follows that the statistic defined by

$$(4.19)\qquad T^2=(\zeta-\mu_4)'\hat{\Sigma}^{-1}(\zeta-\mu_4)/(N-q)\delta^2,$$

has a noncentral Hotelling's -T^2 distribution provided F is also a single vector. Here δ^2 denotes the quadratic form

$$(4.20)\qquad (HH'+HX_0'(XX')^{-1}X_0H' -HX_0(XX')^{-1}F'(F(XX')^{-1}F')^{-1}F(XX')^{-1}X_0H').$$

The above density of T^2 may be used to set confidence regions for the q component vector HX_0'.

Acknowledgement

The author wishes to thank Dr. A. G. Laurent for several helpful suggestions.

WAYNE STATE UNIVERSITY AND KARNATAK UNIVERSITY

REFERENCES

[1] A. Franklin Graybill, *An Introduction to Linear Statistical Models, Vol. 1.* McGraw-Hill, New York, 1961.

[2] D. G. Kabe, "On the generalization of Sverdrup's lemma and its applications to multivariate distribution theory," (Abstract) *Ann. Math. Statist.*, Vol. 31 (1960), p. 817.

[3] A. M. Kshirsagar, "The noncentral multivariate beta distribution," *Ann. Math. Statist.*, Vol. 32 (1961), pp. 104–111.

[4] Erling Sverdrup, "Derivation of the Wishart distribution of the second order sample moments by straightforward integration of a multiple integral," *Skandinavisk Aktuarietidskr.*, Vol. 30 (1947), pp. 151–166.