FOURTH MOMENTS IN THE GENERAL LINEAR MODEL; AND THE VARIANCE OF TRANSLATION INVARIANT QUADRATIC FORMS

by October, 1977

R. D. Anderson\textsuperscript{1/}, R. L. Quaas\textsuperscript{1/} and S. R. Searle\textsuperscript{2/}
Cornell University

Abstract

Using vec and Kronecker product operators, a detailed derivation is given of fourth moments in the general linear model and of the variance of translation invariant quadratic forms.

Introduction

We consider the general linear model

\[ y = X\beta + Z_1u_1 + Z_2u_2 + \cdots + Z_ku_k \]  

where \( X \) of order \( n \times p \) and \( Z_m \) of order \( n \times c_m, m = 1, \ldots, k \) are known incidence matrices, \( \beta \) is an unknown vector of \( p \) fixed effects, and the \( u_m \), of order \( c_m \times 1 \) for \( m = 1, \ldots, k \), are unknown vectors of random effects such that

(i) the elements of \( u_m \) are independent having common variance \( \sigma_m^2 \) and kurtosis \( \gamma_m \), and

(ii) \( u_m \) and \( u_{m'} \) are independent for \( m \neq m' \).

\textsuperscript{1/} Department of Animal Science.
\textsuperscript{2/} Biometrics Unit
\textsuperscript{*} Paper No. BU-630-M in the Biometrics Unit.
Accordingly, the variance-covariance matrix of the vector $\mathbf{y}$ is
\[
\text{Var}(\mathbf{y}) = \frac{k}{m=1} \sigma_{m,m'}^2 Z_{m,m'} = \mathbf{V}.
\]

As far as the ensuing algebra is concerned, a more convenient representation of (1) is
\[
\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} \quad (2)
\]
where $\mathbf{Z}$ is the $n \times c$ partitioned matrix $\mathbf{Z} = [\mathbf{Z}_1 \cdots \mathbf{Z}_k]$, and $\mathbf{u}$ is the $c \times 1$ vector $\mathbf{u}' = [u_1' \cdots u_k']$. Corresponding to model (2), the variance-covariance matrix of $\mathbf{y}$ is then
\[
\mathbf{V} = \mathbf{ZDZ}'
\]
where
\[
\mathbf{D} = \mathbf{E}(\mathbf{u}\mathbf{u}') = \sum_{m=1}^{k} \sigma_{m,m}^2 \mathbf{I}_{c_m}
\]
and $\sum_{\mathbf{A}}$ denotes the direct sum of matrices $\mathbf{A}$.

**Fourth Moments**

The matrix of central fourth moments of the vector $\mathbf{y}$ is, by definition,
\[
\mathbf{F} = \text{Var}[(\mathbf{y} - \mathbf{X}\beta) \ast (\mathbf{y} - \mathbf{X}\beta)] \quad (4)
\]
where $\ast$ is the direct (Kronecker) product of $\mathbf{A}$ and $\mathbf{B}$. Substituting in terms of (2),
\[ F = \text{Var}(Z_u * Z_u), \]
\[ = \text{Var}((Z * Z)(u * u)) \]
\[ = (Z * Z)[\text{Var}(u * u)](Z * Z)' . \]  

(5)

Defining the vec of a matrix to be the \( \vec{M} \) matrix first introduced by Roth [1934], namely, the vector obtained from stacking the columns of the matrix one beneath the other in a single vector, and noting that \( \text{vec}(uu') = u * u \) it follows that

\[ \text{Var}(u * u) = \mathbb{E}[(u * u)(u * u)'] - [\mathbb{E}(u * u)][\mathbb{E}(u * u)]' \]
\[ = \mathbb{E}[(uu') * (uu')] - \mathbb{E}[(uu')][\mathbb{E}(uu')]' \]
\[ = \mathbb{E}[(uu') * (uu')] - \text{vec}(uu')[\text{vec}(uu')]' \]
\[ = \mathbb{E}[(uu') * (uu')] - \text{vec}(D)(\text{vec}D)' , \]  

(6)

on using \( D \) of (3).

To simplify (6) we define

\[ c = c = c_1 + c_2 + \cdots + c_k \]  

(7)

and

\[ \sim w = D \frac{\dot{g}}{\sim} u = \{w_i\} \quad i = 1, \cdots, c . \]  

(8)

Then \( \sim w \) has the properties

\[ \mathbb{E}(\sim w) = 0, \quad \mathbb{E}(\sim w') = \text{var}(\sim w) = I_c \]  

(9)

and

\[ \mathbb{E}(w_i^a) = 3 + \dot{y}_i \quad \text{for } i = 1, \cdots, c \]  

(10)

where
\[
\hat{\gamma}_i = i^{th} \text{ diagonal element of } \sum_{m=1}^{\infty} \gamma_{m-m}^+ \cdot (11)
\]

Then for (6)

\[
E[(u_s u'_s) \ast (u_s u'_s)] = (D^{\delta} \ast D^{\delta})E[(w_s w'_s) \ast (w_s w'_s)](D^{\delta} \ast D^{\delta})',
\]

\[
= (D^{\delta} \ast D^{\delta})E(D^{\delta} \ast D^{\delta}) \quad (12)
\]
on defining

\[
\Sigma_{c^2 \times c^2} = \{\Sigma_{ij}\} \quad \text{for } i, j = 1, 2, \cdots, c
\]

\[
= \{E(w_i w_j w'_s w'_s)\}
\]

\[
= \{E(w_i w_j w'_s w'_s)\} \quad \text{for } i, j, k, \ell = 1, 2, \cdots, c \quad (13)
\]

Now for \(i = j\)

\[
E(w_i w_i w_i w_i) = \begin{cases} 
3 + \hat{\gamma}_i & \text{when } i = k = \ell \\
1 & \text{when } i \neq k = \ell \\
0 & \text{otherwise} 
\end{cases} \quad (14)
\]
and for \(i \neq j\)

\[
E(w_i w_i w_i w_i) = \begin{cases} 
1 & \text{when } i = k, j = \ell \\
1 & \text{when } i = \ell, j = k \\
0 & \text{otherwise}, 
\end{cases} \quad (15)
\]
so that, on defining

\[
e_{s} = i^{th} \text{ column of } I_c, \quad (16)
\]

(14) and (15) give the sub-matrices of (13) as

\[
\Sigma_{ii} = I + (2 + \hat{\gamma}_i)e_i e'_i \quad (17)
\]
and

\[ \Sigma_{ij} = e_i e_j' + e_j e_i' \text{ for } i \neq j. \]  

(18)

Therefore in (13), noting that \( 2e_i e_j' = e_i e_j' + e_j e_i' \text{ for } i = j, \)

\[ \Sigma = \{\Sigma_{ij}\} = I + \{e_i e_j' + e_j e_i'\} \text{ for } i, j = 1, \ldots, c, + \sum_{i=1}^{c} \gamma_i e_i e_i'. \]  

(19)

In (19) it is important to note that for the matrix \[ e_i e_j' + e_j e_i' \text{ the sequence of subscripts is } j = 1, \ldots, c \text{ within } i = 1, \ldots, c. \] This being so, it can be noted that

\[ \{e_i e_i'\} \text{ for } i, j = 1, \ldots, c, = I \) \(c,c,\),

(20)

the permuted identity matrix of order \(c^2\), as used by Tracy and Dwyer [1969] and MacRae [1974]; and

\[ \{e_i e_j'\} \text{ for } i, j = 1, \ldots, c, = \text{vec}(\text{vecI})' \).  

(21)

Furthermore, using (11) in the last term of (19) gives, for \( t_{m-1} = c_1 + c_2 + \cdots + c_{m-1} \)

\[ \sum_{i=1}^{c} \gamma_i e_i e_i' = \sum_{m=1}^{k} \gamma_m \left[ \sum_{i=t_{m-1}+1}^{t_m} e_i e_i' \right]. \]  

(22)

Now for \( i = 1, \ldots, c \)

\( e_i e_i' \) = a diagonal matrix of order \( c \) with its only non-zero element being 1 in the \((i,i)\) position

and

\[
\begin{bmatrix}
  e_i e_i' & 0 \\
  0 & e_{i+1} e_{i+1}'
\end{bmatrix}
\]  

a diagonal matrix of order 2\( c \) with its only non-zero elements being 1 in the \((i,i)\) and \((c+i+1, c+i+1)\) positions.
Consider just the diagonal elements of this last matrix. Between the two 1's there are $c - i + i = c$ zeros; and this is true for all $i$. Denote a row vector of $c$ zeros by $\mathbf{0}_c'$. Then the diagonal elements of (22) are

\[
[y_1 \mathbf{0}_c' y_1 \mathbf{0}_c' \cdots y_1 \mathbf{0}_c' y_2 \mathbf{0}_c' \cdots y_2 \mathbf{0}_c' \cdots y_k \mathbf{0}_c' \cdots y_k \mathbf{0}_c']
\]

(23)

where $y_m$ occurs $c_m$ times for $m = 1, 2, \ldots, k$. This is a vector of $c^2$ elements and by its nature is $\text{vec} \left( \sum_{m=1}^{k} y \mathbf{I}_{m \times c_n} \right)$. Hence, on using the definition

\[
\text{diag} \mathbf{x} = \text{diagonal matrix with diagonal elements being the elements of the vector } \mathbf{x},
\]

we have (22) as

\[
\sum_{i=1}^{c} y_1 \mathbf{e}_i \mathbf{e}_i' = \text{diag} \{ \text{vec} \left( \sum_{m=1}^{k} y \mathbf{I}_{m \times c_n} \right) \}.
\]

(24)

Substituting (20), (21) and (24) into (19) gives

\[
\Sigma = \mathbf{I}_{c} + I_{c, c} + \left( \text{vec} \mathbf{I}_{c} \right) \left( \text{vec} \mathbf{I}_{c} \right) + \text{diag} \{ \text{vec} \left( \sum_{m=1}^{k} y \mathbf{I}_{m \times c_n} \right) \}.
\]

(25)

Using $\Sigma$ in (12) now gives

\[
\mathbb{E} [ (\mathbf{u}_n' \ast \mathbf{u}_n') ] = (\mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c}) \left[ \mathbf{I} + I_{c, c} + \left( \text{vec} \mathbf{I}_{c} \right) \left( \text{vec} \mathbf{I}_{c} \right) + \text{diag} \{ \text{vec} \left( \sum_{m=1}^{k} y \mathbf{I}_{m \times c_n} \right) \} \right] (\mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c})
\]

\[
= \mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c} + \left( \mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c} \right) I_{c, c} \left( \mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c} \right) + \mathbf{z}_n' + \mathbf{z}_n + \Gamma
\]

(26)

where we define

\[
\Gamma = (\mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c}) \left[ \text{diag} \{ \text{vec} \left( \sum_{m=1}^{k} y \mathbf{I}_{m \times c_n} \right) \} \left( \mathbf{D}_{c, c}' \ast \mathbf{D}_{c, c} \right) \right].
\]

(27)
Also, use is made of the results in MacRae [1974] that

\[ I_{(p,p)}(A_{p\times q} \ast B_{p\times q})I_{(q,q)} = B_{p\times q} \ast A_{p\times q} \]  

(28)

and

\[ [I_{(p,p)}]^2 = I_{p^2} \]  

(29)

so that

\[ I_{(p,p)}(A_{p\times q} \ast A_{p\times q}) = (B_{p\times q} \ast A_{p\times q})I_{(q,q)} \]  

(30)

Hence for (26)

\[ (D_1^2 \ast D_2^2)I_{(c,c)}(D_1^2 \ast D_2^2) = (D_1 \ast D_2)I_{(c,c)} \]  

(31)

Furthermore, in (26)

\[ z = (D_1^2 \ast D_2^2)\text{vec}I = \text{vec}D \]  

(32)

because, in general,

\[ \text{vec}(AB\text{c}) = (C' \ast A)\text{vecB} \]  

(33)

as in Neudecker [1969].

Using (31) and (32) in (26) therefore gives

\[ E[(uu') \ast (uu')] = (D_1 \ast D)(I + I_{(c,c)}) + (\text{vecD})(\text{vecD})' + \Gamma \]  

(34)

where \( \Gamma \) is as defined in (27), and so substitution into (6) gives

\[ \text{var}(u \ast u) = (D_1 \ast D)(I + I_{(c,c)}) + \Gamma \]  

(35)

Putting (35) into (5) gives the matrix of fourth moments as
\[ F = (Z \ast Z)[(D \ast D)(I + I(c,c)) + \Gamma](Z' \ast Z') \]

and because \( ZDZ' = \Gamma \) this is:

\[ F = V \ast V + (ZDZ)(I(c,c))(Z' \ast Z') + (Z \ast Z)\Gamma(Z' \ast Z') . \]

Using (30) again leads to

\[ F = (V \ast V)(I + I(n,n)) + (Z \ast Z)\Gamma(Z' \ast Z') \hspace{1cm} (36) \]

and on using (27) this has the equivalent form

\[ F = (V \ast V)(I + I(n,n)) + (ZD^2 \ast ZD^2)[\text{diag}(\text{vec}(\Sigma_k \gamma m\Gamma c_n))](D^2 Z' \ast D^2 Z'). \hspace{1cm} (37) \]

This, then, is the general expression for the matrix of fourth central moments of the vector of observations in linear model theory.

In the special case of normality assumptions, i.e., \( u_m \sim N(0, \sigma^2 \Gamma c_n) \), we have \( \gamma_m = 0 \) and (37) reduces to

\[ F = (V \ast V)(I + I(n,n)) \hspace{1cm} (38) \]

**Variance of Translation Invariant Quadratic Forms**

The quadratic form \( y'\tilde{\tilde{y}} \) is called translation invariant when \( A \), as well as being symmetric, satisfies \( AX = 0 \). Then the variance of the translation invariant quadratic form is

\[
v(y'\tilde{\tilde{y}}) = v[(y - X\tilde{\tilde{y}})'A(y - X\tilde{\tilde{y}})] \\
= v(u'ZAZu) \\
= v[\text{tr}[A(Zuu'Z')]].
\]
Now use the general result for any product \( P_{\sim}Q_{\sim} \), that

\[
\text{tr}(P_{\sim}Q_{\sim}) = (\text{vec}_{\sim}P_{\sim})'\text{vec}_{\sim}Q_{\sim}
\]

and so

\[
\nu(y' Ay) = \nu((\text{vec}_{\sim}A)'\text{vec}(Z_{uu}'Z_{\sim}'))
\]

\[
= (\text{vec}_{\sim}A)'\text{var}[\text{vec}(Z_{uu}'Z_{\sim}')]\text{vec}_{\sim}A
\]

\[
= (\text{vec}_{\sim}A)'\text{var}[(Z_{\sim})' (Z_{\sim})]'\text{vec}_{\sim}A
\]

\[
= (\text{vec}_{\sim}A)'F(\text{vec}_{\sim}A), \text{ using (4)}
\]

and on using (36) this gives

\[
\nu(y' Ay) = \theta_1 + \theta_2
\]

for

\[
\theta_1 = (\text{vec}_{\sim}A)'(V * V)(I + I_{(n,n)})\text{vec}_{\sim}A
\]

and

\[
\theta_2 = (\text{vec}_{\sim}A)'(ZD_{\sim}^{\frac{1}{2}} * ZD_{\sim}^{\frac{1}{2}})[\text{diag}\{\text{vec}(\sum_{m=1}^{k} \gamma_m I_{m-c_{\sim}})\}](D_{\sim}^{\frac{1}{2}}Z' * D_{\sim}^{\frac{1}{2}})'\text{vec}_{\sim}A.
\]

In \( \theta_1 \) of (42), the elements of the \([(i-1)n+j]^\text{th} \) row of \( \sim_{(n,n)} \) are all zero except for a 1 in the \([(j-1)n+i]^\text{th} \) column (and vice versa); and also, because \( \sim_{\sim} \) is symmetric, the \([(i-1)n+j]^\text{th} \) and \([(j-1)n+i]^\text{th} \) elements of \( \text{vec}_{\sim}A \) are the same.

Hence

\[
I_{(n,n)}(\text{vec}_{\sim}A) = \text{vec}_{\sim}A.
\]

Also using (33) and (39)

\[
(\text{vec}_{\sim}A)'(V * V)\text{vec}_{\sim}A = (\text{vec}_{\sim}A)'\text{vec}(VAV) = \text{tr}(AV)^2,
\]
so that

$$\theta_1 = 2\text{tr}(AV)^2.$$  \hspace{1cm} (46)

Simplification of $\theta_2$ in (43) starts with using (33) to get

$$\theta_2 = \left[\text{vec}\left(\frac{1}{2}Z'AZ^2\right)\right]'\left[\text{diag}\left[\text{vec}\left(\sum_{m=1}^{k} \frac{1}{m-c_m}\right)\right]\right]\text{vec}\left(\frac{1}{2}Z'AZ^2\right)$$

which is of the form

$$\theta_2 = \left(\text{vec}\mathcal{H}\right)'\left[\text{diag}[\text{vec}\mathcal{L}]\right]\text{vec}\mathcal{H}$$ \hspace{1cm} (47)

for

$$\mathcal{H} = \frac{1}{2}Z'AZ^2 \text{ and } \mathcal{L} = \sum_{m=1}^{k} \frac{1}{m-c_m}. \hspace{1cm} (48)$$

The nature of the vec and diag operators means that (47) is

$$\theta_2 = \sum_{i,j} \Sigma h^2_{ij} \delta_{ij}$$ \hspace{1cm} (49)

for $\mathcal{H} = \{h_{ij}\}$ and $\mathcal{L} = \{L_{ij}\}$ of (48). But with this $\mathcal{L}$, the only non-zero $L_{ij}$'s are the diagonal ones, $L_{tt} = \gamma_m$ for $t = 1, \cdots, c_m$ and $m = 1, \cdots, k$. Furthermore, as in (23), these diagonal elements have $c$ zeros between them in vec$\mathcal{L}$ so that the use of (48) in (49) gives

$$\theta_2 = \sum_{m=1}^{k} \sum_{t=1}^{c_m} \Sigma \Sigma h^2_{tt}$$

for

$$h^2_{tt} = t^{th} \text{ diagonal element of the } m^{th} \text{ diagonal sub-matrix of } \frac{1}{2}Z'AZ^2 \text{ and }$$

$$\sum_{m=1}^{k} \sum_{t=1}^{c_m} h^2_{tt} = \sigma^2_{m}(t^{th} \text{ diagonal element of the } m^{th} \text{ diagonal sub-matrix of } Z'AZ).$$
Therefore

\[ \theta_2 = \sum_{m=1}^{k} \gamma_m c_m^2 (\text{sum of squares of diagonal elements of } Z'AZ_{m\sim m}). \quad (50) \]

Substituting (46) and (50) into (41) gives

\[ v(y'Ay) = 2tr(AV)^2 + \sum_{m=1}^{k} \gamma_m c_m^2 (\text{sum of squares of diagonal elements of } Z'AZ_{m\sim m}). \quad (51) \]

This is the variance, under non-normality, of a translation invariant \((AX = 0)\) quadratic form \(y'Ay\). Under normality, \(\gamma_m = 0\) for all \(m\) and (51) reduces to the familiar form

\[ v(y'Ay) = 2tr(AV)^2. \quad (52) \]

Equation (51) is, of course, equivalent to the result given by Rao [1971] where he writes \(\Delta_1^\sim\) for \(D\) and \(\Delta_2^\sim = \sum_{m=1}^{k} \gamma_m c_m^2 \mathbf{1} \mathbf{1}'\) and \(\mathbf{B} = \text{diag}\{\text{diagonal elements of } B\}\) and so gets

\[ v(y'Ay) = 2tr(B\Delta_1^\sim)^2 + tr(\mathbf{B}\Delta_2^\sim \mathbf{B}), \]

for \(B = Z'AZ\). With \(V\) being \(ZDZ'\) this is readily seen to be the same as (51).

References


