UNCORRELATED LINEAR TRANSFORMATIONS OF RESIDUALS IN
MULTIPLE REGRESSION

BU-216-M D. S. Robson April, 1966

Abstract

If $G$ is a generalized inverse of $X'X = (X_1'X_1 + X_2'X_2)$ and $\&$ is a generalized inverse of $X_1'X_1$ then

$$Y_1 - X_1\&X_1'Y_1 = Y_1 - X_1GX_1'Y + X_1GX_2'K(Y_2 - X_2GX_1'Y)$$

where $K$ is a generalized inverse of $I - X_2GX_2'$. Further, if $Y = X\beta + \epsilon$ and $E(\epsilon\epsilon') = \sigma^2I$ then $E(Y_1 - X_1\&X_1'Y)(Y_2 - X_2GX_1'Y) = \phi$.

Biometrics Unit, Plant Breeding Department, Cornell University.
In an earlier note (BU-210-M) it was shown that if the regression model

\[
Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \beta + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}, \quad E(\epsilon\epsilon') = \sigma^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}
\]

is fitted separately to the first \( k \) observations \( Y_1 \) and to all \( n \) observations \( Y \) then the least squares residuals in the first instance

\[
f_1 = Y_1 - X_1'X_1'Y_1, \quad X'X_1'X_1 = X_1'X_1
\]

are uncorrelated with the \( n-k \) least squares residuals

\[
e_2 = Y_2 - X_2'X_1'Y_1, \quad X'X_2'X_2 = X_2'X_2
\]

obtained after fitting the model to all \( n \) observations. The purpose of the present note is to show that

\[
f_1 = e_1 + X_1'X_2'Ke_2
\]

for all \( K \) such that

\[
(I - X_2'X_2')K(I - X_2'X_2') = I - X_2'X_2'
\]
Rewriting (1) in terms of \( Y_1 \) and \( Y_2 \) we get

\[
e_1 + X_1 G X_2^T K e_2 = (I - X_1 G X_1^T) Y_1 - X_1 G X_2^T Y_2
\]

\[
+ X_1 G X_2^T K (I - X_2 G X_2^T) Y_2 - X_1 G X_2^T K X_2 G X_1^T Y_1
\]

The coefficient of \( Y_2 \) in (3) is null since

\[
[I - X_2 G X_2^T] K [X_1 G X_1^T K (I - X_2 G X_2^T) - X_1 G X_1^T]
\]

\[
= (I - X_2 G X_2^T) K' X_2 G' X_1^T X_1 G X_2^T K (I - X_2 G X_2^T) - X_2 G X_1^T K (I - X_2 G X_2^T)
\]

\[
- (I - X_2 G X_2^T) K' X_2 G' X_1^T X_1 G X_2^T + X_2 G' X_1^T X_1 G X_2^T
\]

and from the relation \( X G X' X = X \) we have

\[
X_2 G' X_1^T X_1 = (I - X_2 G' X_2^T) X_2
\]

and from the invariance of \( X G X' \) then

\[
X_2 G' X_1^T X_1 = (I - X_2 G X_2^T) X_2
\]

Upon substituting this expression into (4) and utilizing the relation (2) we see that the right hand side of (4) is null, and hence

\[
X_1 G X_2^T K (I - X_2 G X_2^T) = X_1 G X_2^T
\]
Equation (3) thus reduces to

\[ e_1 + x_1^2 k e_2 = (I - x_1^2 (G + G x_2^2 G) x_1^2) y_1 \]

and since

\[ f_1 = (I - x_1^2 x_1^2) y_1 \]

then to establish (1) we must show that \( G + G x_2^2 G \) is a generalized inverse of \( X_1^2 X_1 \); i.e., that

\[ X_1^2 X_1 (G + G x_2^2 G) X_1^2 X_1 = X_1^2 X_1 \]

As in (5) we may write

\[ X_1^2 X_1 (G + G x_2^2 G) X_1^2 X_1 = X_2 X_1 - X_2 X_2 G X_1 X_2 \]

so that (6) reduces to

\[ X_1^2 X_1 (G + G x_2^2 G) X_1^2 X_1 = X_2 X_2 G X_1 X_2 \]

From (5), the left hand side of (7) may be written as

\[ x_2^2 (I - x_2^2 G x_2^2) = x_2^2 (I - x_2^2 G x_2^2) x_2 \]
and applying (5) to the right hand side of (7) gives the same result, thus confirming (6) and hence also (1).