THE REDUCTION IN SUM OF SQUARES ATTRIBUTABLE TO A SUBSET OF A SET OF REGRESSION COEFFICIENTS AND THE INVARIANCE UNDER CERTAIN LINEAR TRANSFORMATIONS OF A SEQUENCE OF QUADRATIC FORMS IN THESE COEFFICIENTS

BY

Bruce McConnell Graham

A Thesis Submitted to the Graduate Committee

for the Degree of

MASTER OF SCIENCE

in

Statistics

Approved:

_____

Course Adviser

_____

Head of Department

_____

Dean of Agriculture

Virginia Polytechnic Institute

1947

152490

TABLE OF CONTENTS

	Page
Statement of the Problem	1
Special Cases and Applications	6
A General Solution	13
Conclusions	22
Bibliography	23

ACKNOWLEDGEMENT

The writer is very grateful for the guidance and constructive criticism which he has received from Dr. D. B. DeLury, under whose supervision this paper was prepared.

I Statement of the Problem

The general problem of linear regression may be described as follows. Let y_α be the observed value of the dependent variable and $x_{0\alpha}, x_{1\alpha}, \ldots, x_{q\alpha}, \ldots, x_{p\alpha}$ the observed values of the independent variables which are linearly related to y_α. $\beta_0, \beta_1, \ldots, \beta_q, \ldots \beta_p$ are unknown population parameters whose values are to be estimated from the relations

$$(1) \quad y_\alpha = \beta_0 x_{0\alpha} + \beta_1 x_{1\alpha} + \ldots + \beta_q x_{q\alpha} + \ldots + \beta_p x_{p\alpha} + \epsilon_\alpha .$$

The x's are known, and the ϵ's are usually assumed to be normally and independently distributed. The β's are estimated by the method of least squares; that is, estimates b_i of the population parameters β_i are found such that the sum of the squared deviations of the estimated from the observed y's is as small as possible. The regression equation takes the form

$$(2) \quad Y_\alpha = b_0 x_{0\alpha} + b_1 x_{1\alpha} + \ldots + b_p x_{p\alpha},$$

where the b_i are calculated so that

$$(3) \quad \sum_{\alpha=1}^{n} (y_\alpha - Y_\alpha)^2 = \text{minimum.}$$

When the expression (3) is minimized, we obtain a set of normal equations of the form shown below:

$$(4) \quad \begin{array}{l}
b_0 S(x_0^2) + b_1 S(x_0 x_1) + b_2 S(x_0 x_2) \ldots + b_p S(x_0 x_p) = S(x_0 y) \\
b_0 S(x_1 x_0) + b_1 S(x_1^2) + b_2 S(x_1 x_2) \ldots + b_p S(x_1 x_p) = S(x_1 y) \\
b_0 S(x_2 x_0) + b_1 S(x_2 x_1) + b_2 S(x_2^2) \ldots + b_p S(x_2 x_p) = S(x_2 y) \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
b_0 S(x_p x_0) + b_1 S(x_p x_1) + b_2 S(x_p x_2) \ldots + b_p S(x_p^2) = S(x_p y).
\end{array}$$

In a more condensed notation these same normal equations may be written as

$$(5) \quad \sum_{j=0}^{p} a_{ij}b_j = g_i \quad , \quad i=0,1,2,\ldots,p.$$

Here $a_{ij} = S(x_i x_j)$ and $g_i = S(x_i y)$, where S denotes summation over the entire sample.

The sum of the squares of all the y's will be called the total sum of squares and may be written Sy^2. The sum of squares $S(y-Y)^2$ is called the residual sum of squares, and the difference between the total and the residual sum of squares is referred to as the reduction in sum of squares attributable to the regression coefficients b_i. With the present notation this latter sum may be written

$$(6) \quad \sum_{i=0}^{p} b_i g_i \ .$$

In some cases it is convenient to assign to x_0 the constant value one; this convention will be adopted in the following discussion.

If we make a slight modification and measure the variables from their means, all the terms in the first row and the first column of the normal equations shown above will vanish, and the normal equations may now be expressed

$$(7) \quad \sum_{j=1}^{p} a_{ij}b_j = g_i \quad , \quad i=1,2,\ldots,p,$$

and the reduction in sum of squares due to a regression on the set of x's $x_1, x_2, \ldots, x_p$ becomes

$$(8) \quad \sum_{i=1}^{p} b_i g_i .$$

This is a well known special case of the problem which is to be solved

for the general case. The values of $b_1, b_2, \ldots, b_p$ are the same whether b_0 is included or not, but the g_i have different values in the two cases. We will work with the equations in form (7) because this is more convenient.

This reduction in sum of squares can also be written in two other forms,

$$(9) \quad \sum_{i,j=1}^{p} a_{ij} b_i b_j \quad \text{and}$$

$$(10) \quad \sum_{i,j=1}^{p} c_{ij} g_i g_j \quad \text{, where } c_{ij} \text{ is defined such that}$$

$$\sum_{j=1}^{p} a_{ij} c_{jk} = \delta_{ik} \quad \text{, and } \delta_{ik} \text{ is the Kronecker delta.}$$

After the computations have been carried out and the reduction in sum of squares calculated, the question often arises as to the part of this reduction which may be associated with a regression on a specified subset of the independent variables.

The direct method for computing the reduction in sum of squares due to the subset $x_{q+1}, \ldots, x_p$ runs as follows:

First fit the regression equation

$$Y = b_1 x_1 + b_2 x_2 + \ldots + b_q x_q + b_{q+1} x_{q+1} \ldots + b_p x_p .$$

The reduction in sum of squares due to this fitting is given by

$b_1 g_1 + b_2 g_2 + \ldots + b_q g_q + b_{q+1} g_{q+1} \ldots + b_p g_p$, the residual sum of squares in this case being

$$S(y-Y)^2 = \sum y^2 - b_1 g_1 - b_2 g_2 - \ldots - b_q g_q - \ldots b_p g_p .$$

If we then fit to only the first q independent variables

$$Y' = b'_1x_1 + b'_2x_2 + \dots b'_qx_q \ .$$

The reduction in sum of squares due to this second fitting is

$b'_1g_1 + b'_2g_2 + \dots + b'_qg_q$ and the residual sum of squares

$$S(y-Y')^2 = \sum y^2 - b'_1g_1 - b'_2g_2 - \dots - b'_qg_q \ .$$

(11) $S(y-Y')^2 - S(y-Y)^2 \equiv S(Y-Y')^2$ gives the reduction in sum of squares due to a regression on the variables $x_{q+1}, x_{q+2}, \dots x_p$.

There are numerous cases in which it would be convenient to have a different method than the direct approach outlined above. Some alternative methods applicable to a restricted set of cases have been proposed. It is the purpose of this paper to present a general solution for the reduction in sum of squares (11).

It will be shown in this paper

$$(12) \quad S(Y-Y')^2 = \sum_{i,j=q+1}^{p} a^*_{ij}b_ib_j$$

where the a^*_{ij} are values which are obtained at a certain stage of the solution of the normal equations when the reduction is carried out according to Chio's rule; incidentally, this is the method which must be followed to obtain valid results with Nair's formula.[1] The method here shown is applicable in the general case of least squares and includes Nair's result as a special case.

As will be shown in the schematic solution in section II, where the case p = 4 is used as an illustration,

$$\sum_{j=q+1}^{p} a^*_{ij}b_j = g^*_i \ , \quad i=q+1,\dots,p, \text{ where } q = 1, 2, 3 \ .$$

For example, when q=2, these equations represent the normal equations at the second stage of the reduction,

(1) See page 8.

$$a_{33.2}b_3 + a_{34.2}b_4 = g_{3.2} ,$$

$$a_{43.2}b_3 + a_{44.2}b_4 = g_{4.2} ,$$

which may be written in the form

$$\sum_{j=3}^{4} a^*_{ij}b_i = g^*_i , \quad i = 3,4 .$$

Therefore, by an appropriate substitution, (12) may be written in the equivalent form

$$(13) \quad \sum_{i=q+1}^{p} b_i g^*_i .$$

The solution of the reduced set of normal equations at any stage may be written

$$b_i = \sum_{i=q+1}^{p} c_{ij} g^*_i , \quad j=q+1,\ldots,p.$$

Now by a substitution in (13) we obtain another equivalent form

$$(14) \quad \sum_{i,j=q+1}^{p} c_{ij}\, g^*_i g^*_j .$$

Thus we will have three alternatives to the general approach which may prove more convenient to use. Equations (8), (9), and (10) are special forms corresponding to the three general solutions.

II Special Cases and Applications

One special case of the problem covets the reduction in sum of squares due to b_i alone. This b_i^2/c_{ii}, where b_i and c_{ii} are defined as in the preceding section. This result is already well known.

Another special case, also well known, in which formula (12) is easily shown to hold is the case in which the independent variables are transformed, so that the regression equation is fitted to linear functions of the independent variables which are orthogonal over the set of observed values. If we make the transformation

$\xi_i = \sum_{k=1}^{i} \lambda_{ik} x_k$, $i=1,2,\ldots,p$, choosing the λ_{ik} so that

$S\, \xi_i\, \xi_j = 0$, $i \neq j$, then the two regression equations fitted in the direct method of solution are, in terms of the ξ's ,

$$Y = B_1 \xi_1 + B_2 \xi_2 + \ldots + B_q \xi_q + B_{q+1} \xi_{q+1} + \ldots + B_p \xi_p \quad \text{and}$$

$$Y' = B_1 \xi_1 + B_2 \xi_2 + \ldots + B_q \xi_q \ .$$

The normal equations determining the B's are

$B_i S\, \xi_i^2 = Sy\, \xi_i$, and the elements of the inverse matrices are

$c_{ij} = \delta_{ij} / S\, \xi_i^2$, $i,j = 1,2,\ldots,q$, $q+1,\ldots,p$.

The residual sums of squares are

$$S(y-Y)^2 = Sy^2 - \sum_{i=1}^{p} B_i\,(Sy\,\xi_i) = Sy^2 - \sum_{i=1}^{p} B_i^2/C_{ii} \text{ and}$$

$$S(y-Y')^2 = Sy^2 - \sum_{i=1}^{q} B_i\,(Sy\,\xi_i) = Sy^2 - \sum_{i=1}^{q} B_i^2/C_{ii}.$$

The reduction in sum of squares (12) thus takes the form

$$S(Y-Y')^2 = \sum_{i=q+1}^{p} B_i^2/C_{ii}.$$

This is the form taken by formula (12) in this special case. When it is shown that this sum of squares is invariant under the transformation that put it in this form, then the truth of the formula is established. This invariance is proved in Section III.

For an instance in which these formulas are needed, we will consider the case of a multiple classification, which is defined as a simultaneous classification of several different sets of classes. For instance, a randomized blocks field plot arrangement (in which each treatment appears at a random position in each block) is an example of a double classification; each y observation is classified as to block and treatment. If y_{ij} is the yield of the plot in the ith block which receives the jth treatment, the following hypothesis is set up:

$$y_{ij} = \mu + \beta_i + \tau_j + \epsilon_{ij},$$

where μ is the true mean yield of the experiment, β_i is the effect of the ith block and τ_j the effect of the jth treatment. These are the population parameters to be estimated, and ϵ_{ij} is the experimental error of the plot. It is assumed that the various β_i and the τ_j exert their influence independently of one another. That is to

say, referring to the randomized blocks layout, individuals belonging to the ith block will receive a constant effect β_i whatever be the subclass of τ to which they belong, and vice versa.

In the randomized blocks arrangement the blocks are ordinarily orthogonal with the treatments. However, it may happen that several plot yields are missing, which is a more general class of a double classification. In the randomized blocks design x_i has the value 1 for every y_{ij} in block i and the value zero elsewhere; also x_j has the value 1 for every y_{ij} in treatment j and the value zero elsewhere.

Note that whether the data are orthogonal or not, the x's are in every case assigned certain formal values, which makes them in fact only dummy variables. In the general case under discussion there are no restrictions on the meaning which the x's can have.

One formula for obtaining the reduction in sum of squares (11) without explicitly fitting a second regression equation is given in a paper by K. R. Nair. This paper, however, applies only to the special case of a double classification as described above; it is shown that in this instance the reduction in sum of squares (11) is given by the formula

$$\sum_{i=q+1}^{p} b_i o_i ,$$

where the o_i are the right hand sides of the normal equations which remain after $b_1 \ldots b_q$ have been eliminated. Nair does not discuss the point that the elimination must be carried out according to an acceptable pattern before this equality can hold.

A schematic layout for the solution of the b_i , the c_{ij} , the a^*_{ij}, and the g^*_i as defined in this paper will now be given. This

layout will be given for the specific case $p = 4$, but it embodies all the properties of the general case. We have, in fact, a solution of two sets of equations, one being the normal equations

$$\sum_{j=1}^{4} a_{ij}b_j = g_i \ , \quad i = 1,2,3,4, \text{ and the other being}$$

$$\sum_{j=1}^{4} a_{ij}c_{jk} = \delta_{ik} \ , \quad i,k = 1,2,3,4.$$

The regression equation in this case is

$y = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$, the variables being measured from their sample means. The original set of normal equations is first written down, and the four stages in the reduction are then carried out successively. The $a_{ij.k}$ and the $g_{i.k}$ that appear at the kth stage of the reduction are the a^*_{ij} and the g^*_i appropriate when $q = k$ for computing the reduction in sum of squares associated with a regression on the subset $x_{q+1}, x_{q+2}, \ldots, x_p$. The asterisks stand for numbers which are needed in the computation, but whose values are not used in this discussion.

The operations carried out as shown will produce at the kth stage of the reduction the values of $b_{i.k}$ and $c_{ij.k}$ appropriate when $q = k$ for fitting a regression of y on the subset $x_1, x_2, \ldots x_q$. The $b_{i.4}$ and the $c_{ij.4}$ are written simply as b_i and c_{ij} .

The mechanics of the Chio procedure may be described as follows. At the first stage we make these operations: division of the first normal equation by a_{11} to obtain the new first equation; multiplication of the new first equation by a_{21} and subtraction from the second original equation; similar operations on the remaining original equations.

For the next stage, the second equation in the first stage of the reduction is divided by $a_{22.1}$; the other equations in the first stage are operated on as in the first reduction.

This process is continued until all the a_{ij} are replaced by zero's and 1's. The numbers appearing in the column which consisted of 1's in the original normal equations are now the desired regression coefficients b_i . The array on the right hand side now consists of numbers c_{ij} which satisfy the relation

$$\sum_{j=1}^{4} a_{ij}c_{jk} = \delta_{ik} , \quad i,k = 1,2,3,4.$$

In the language of matrix algebra the c_{ij} are elements of the matrix inverse to the coefficient matrix with elements a_{ij}. The Chio method of reduction preserves the elements c_{ij} of the inverse matrix; for example, the elements c_{ij} inverse to the block

$$\begin{matrix} a_{22.1} & a_{23.1} & a_{24.1} \\ a_{32.1} & a_{33.1} & a_{34.1} \\ a_{42.1} & a_{43.1} & a_{44.1} \end{matrix}$$

are the same as they were in the case of

$$\begin{matrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{matrix} .$$

a_{11}	a_{12}	a_{13}	a_{14}	g_1	1	0	0	0
a_{21}	a_{22}	a_{23}	a_{24}	g_2	0	1	0	0
a_{31}	a_{32}	a_{33}	a_{34}	g_3	0	0	1	0
a_{41}	a_{42}	a_{43}	a_{44}	g_4	0	0	0	1

1	*	*	*	$b_{1.1}$	$c_{11.1}$	0	0	0
0	$a_{22.1}$	$a_{23.1}$	$a_{24.1}$	$g_{2.1}$	*	1	0	0
0	$a_{32.1}$	$a_{33.1}$	$a_{34.1}$	$g_{3.1}$	*	0	1	0
0	$a_{42.1}$	$a_{43.1}$	$a_{44.1}$	$g_{4.1}$	*	0	0	1

1	0	*	*	$b_{1.2}$	$c_{11.2}$	$c_{12.2}$	0	0
0	1	*	*	$b_{2.2}$	$c_{21.2}$	$c_{22.2}$	0	0
0	0	$a_{33.2}$	$a_{34.2}$	$g_{3.2}$	*	*	1	0
0	0	$a_{43.2}$	$a_{44.2}$	$g_{4.2}$	*	*	0	1

1	0	0	*	$b_{1.3}$	$c_{11.3}$	$c_{12.3}$	$c_{13.3}$	0
0	1	0	*	$b_{2.3}$	$c_{21.3}$	$c_{22.3}$	$c_{23.3}$	0
0	0	1	*	$b_{3.3}$	$c_{31.3}$	$c_{32.3}$	$c_{33.3}$	0
0	0	0	$a_{44.3}$	$g_{4.3}$	*	*	*	1

1	0	0	0	b_1	c_{11}	c_{12}	c_{13}	c_{14}
0	1	0	0	b_2	c_{21}	c_{22}	c_{23}	c_{24}
0	0	1	0	b_3	c_{31}	c_{32}	c_{33}	c_{34}
0	0	0	1	b_4	c_{41}	c_{42}	c_{43}	c_{44}

In the layout shown the reduction is carried out strictly according to Chio's rule; that is

$$a_{ij.k} = a_{ij.(k-1)} - a_{ik.(k-1)}\, a_{kj.(k-1)} / a_{kk.(k-1)} \text{ and}$$

$$g_{i.k} = g_{i.(k-1)} - a_{ik.(k-1)}\, g_{k.(k-1)} / a_{kk.(k-1)} .$$

There are other methods available which would shorten this calculation somewhat and make the algebra less unwieldy, and if a solution for the b_i only is required some more compact method should undoubtedly be used. As has been shown, however, the Chio procedure generates some numbers as by-products which are useful in further analysis of the data.

If the procedure used in the reduction differs from the Chio method only in the divisors and multipliers, multiples of the a^*_{ij} and the g^*_i, instead of these numbers themselves, are usually obtained at the various stages of the solution. This situation can be recognized by observing in each $g_{i.k}$ the coefficient of g_i. Under the Chio rule this coefficient must be equal to one, and when this is not the case the proper values may be obtained at any stage of the solution by a single multiplication.

III A General Solution

Using the same set of y's as given in the general case described on page one, a general linear transformation may be made on the x's such that

(15) $\xi_i = \sum_{j=1}^{p} \lambda_{ij} x_j$, $i=1, 2, \ldots, p$. Also define τ_{jk} such that

(16) $\sum_{j=1}^{p} \lambda_{ij} \tau_{jk} = \delta_{ik}$, $i,k = 1,2, \ldots, p$.

This transformation will give a new set of regression coefficients which will be called B_i. Now

$$y = b_1 x_1 + b_2 x_2 + \ldots + b_p x_p \equiv B_1 \xi_1 + B_2 \xi_2 + \ldots + B_p \xi_p.$$

The mathematical relations between the original and the transformed systems are given below in detail.

A_{ij} may be expressed in terms of a_{ij} by the substitutions shown.

(17) $$A_{ij} = \mathop{S}_{\alpha=1}^{n} \xi_{i\alpha} \xi_{j\alpha} = \mathop{S}_{\alpha=1}^{n} \sum_{k,m=1}^{p} \lambda_{ik} \lambda_{jm} x_{k\alpha} x_{m\alpha}$$

$$= \sum_{k,m=1}^{p} \lambda_{ik} \lambda_{jm} a_{km} .$$

As the first step in the transition from A_{ij} to a_{ij} we multiply (17) by τ_{ni} and sum over i, obtaining

$$\sum_{i,k,m,=1}^{p} \tau_{ni} \lambda_{ik} \lambda_{jm} a_{km} = \sum_{i=1}^{p} \tau_{ni} A_{ij} .$$

Summing the left side over i and then over k, we have

$$\sum_{k,m=1}^{p} \delta_{kn} \lambda_{jm} a_{km} = \sum_{m=1}^{p} \lambda_{jm} a_{nm} ,$$

and multiplying by $\mathcal{C}_{rj}$ we obtain

$$\sum_{j,m=1}^{p} \lambda_{jm} \mathcal{C}_{rj} a_{nm} = \sum_{i,j=1}^{p} \mathcal{C}_{ni} \mathcal{C}_{rj} A_{ij} .$$

Summing the left side over j and then over m,

$$\sum_{m=1}^{p} \delta_{mr} a_{nm} = a_{nr} ,$$

and we have the relationship

$$(18) \quad a_{nr} = \sum_{i,j=1}^{p} \mathcal{C}_{ni} \mathcal{C}_{rj} A_{ij} .$$

To express C_{ij} in terms of c_{ij}, we proceed as follows. Multiplying (17) by C_{jm} and summing over j we obtain

$$\sum_{j,k,r=1}^{p} \lambda_{ik} \lambda_{jr} a_{kr} C_{jm} = \sum_{j=1}^{p} A_{ij} C_{jm} = \delta_{im} .$$

Multiplying these equations by $\mathcal{C}_{ni}$ and summing gives

$$\sum_{i,j,k,r=1}^{p} \lambda_{ik} \mathcal{C}_{ni} \lambda_{jr} a_{kr} C_{jm} = \sum_{i=1}^{p} \mathcal{C}_{ni} \delta_{im} ,$$

and summing over i and k we obtain

$$\sum_{j,r=1}^{p} \lambda_{jr} a_{nr} C_{jm} = \mathcal{C}_{nm} .$$

Now, multiplying by c_{sn} and summing we have

$$\sum_{j,r,n=1}^{p} \lambda_{jr} c_{sn} a_{nr} C_{jm} = \sum_{n=1}^{p} c_{sn} \mathcal{C}_{nm}$$

$$= \sum_{j,r=1}^{p} \lambda_{jr} \delta_{sr} C_{jm} = \sum_{j=1}^{p} \lambda_{js} C_{jm} .$$

Lastly, multiplying by λ_{mv} and summing, we obtain

$$(19)\quad \sum_{j,m=1}^{p} \lambda_{js}\lambda_{mv}C_{jm} = \sum_{n,m=1}^{p} c_{sn}\lambda_{mv}\mathcal{C}_{nm}$$

$$= \sum_{n=1}^{p} c_{sn}\delta_{vn}$$

$$= c_{sv}.$$

Having obtained this relationship, we now examine its inverse, the expression of c_{ij} as a function of C_{ij}.

Multiplying (19) by $\mathcal{C}_{si}$ and summing we obtain

$$\sum_{s=1}^{p} c_{sv}\mathcal{C}_{si} = \sum_{s,j,m=1}^{p} \lambda_{js}\mathcal{C}_{si}\lambda_{mv}C_{jm}$$

$$= \sum_{j,m=1}^{p} \delta_{ji}\lambda_{mv}C_{jm}$$

$$= \sum_{m=1}^{p} \lambda_{mv}C_{im} .$$

Multiplying these equations by $\mathcal{C}_{vk}$ and summing we have

$$(20)\quad \sum_{v,s=1}^{p} c_{sv}\mathcal{C}_{si}\mathcal{C}_{vk} = \sum_{m,v=1}^{p} \lambda_{mv}\mathcal{C}_{vk}C_{im}$$

$$= \sum_{m=1}^{p} \delta_{mk}C_{im}$$

$$= C_{ik} .$$

G_i is easily written in terms of g_i by substituting $\sum_{k=1}^{p} \lambda_{ik} x_k$ for ξ_i.

$$(21) \quad G_i = \sum_{\alpha=1}^{n} \xi_{i\alpha} y_\alpha = \sum_{k=1}^{p} \lambda_{ik} g_k .$$

Now the inverse relations between g_i and G_i will be shown.

Multiplying (21) by μ_{ri} and summing, we obtain

$$(22) \quad \sum_{i=1}^{p} \mu_{ri} G_i = \sum_{k,i=1}^{p} \mu_{ri} \lambda_{ik} g_k$$

$$= \sum_{k=1}^{p} \delta_{rk} g_k$$

$$= g_r .$$

A solution of the normal equations[7] gives

$$b_r = \sum_{n=1}^{p} c_{rn} g_n ,$$ which, with the transformed values of c_{rn} and g_n

$$= \sum_{i,n,j,m=1}^{p} \lambda_{jr} \lambda_{mn} C_{jm} \mu_{ni} G_i .$$

$$= \sum_{i,j,m=1}^{p} \lambda_{jr} C_{jm} G_i \delta_{im}$$

$$= \sum_{i,j=1}^{p} \lambda_{jr} C_{ji} G_i .$$

But $\sum_{j=1}^{p} C_{ji} G_i = B_i$, and therefore

$$(23) \quad b_r = \sum_{j=1}^{p} \lambda_{jr} B_j .$$

To reverse this process we multiply (23) by μ_{rn} and sum, obtaining

$$(24)\sum_{r=1}^{p}\mu_{rn}b_r = \sum_{j,r=1}^{p}\mu_{rn}\lambda_{jr}B_j$$

$$=\sum_{j=1}^{p}\delta_{nj}B_j$$

$$= B_n.$$

It is shown in the following paragraph that the sum of squares (13) is invariant under this transformation. Thus, the sum of squares associated with a regression on $x_1, x_2, \ldots, x_p$ is the same as that associated with a regression on $\xi_1, \xi_2, \ldots, \xi_p$.

$$\sum_{r,n=1}^{p} a_{nr}b_nb_r = \sum_{r,n=1}^{p}\ \sum_{i,j,k,m=1}^{p}\mu_{ni}\mu_{rj}A_{ij}\lambda_{km}B_k\lambda_{mr}B_m$$

$$=\sum_{i,j,k,m=1}^{p}\delta_{ik}\delta_{jm}A_{ij}B_kB_m$$

$$=\sum_{k,m=1}^{p}A_{km}B_kB_m .$$

In order to study the relationship between the original and the transformed systems, considering only the subset $x_{q+1}, x_{q+2}, \ldots, x_p$, the restriction $\lambda_{ij}=0, j>i$, is put on the transformation

$\xi_i = \sum_{j=1}^{p}\lambda_{ij}x_j$. If $\lambda_{ij}= 0, j>i$, it follows from the equations

$$\sum_{j=1}^{p}\lambda_{ij}\mu_{jk} = \delta_{ik}$$

that $\mathcal{C}_{jk} = 0$, $k>j$. Under this restriction it is now easily shown that the b_r and the c_{rn} satisfy the same relationships over the restricted range corresponding to $x_{q+1}, \ldots, x_p$ that held over the range $x_1, \ldots, x_p$. By equation (23)

$$b_r = \sum_{j=1}^{p} \lambda_{jr} B_j .$$

But $\lambda_{jr} = 0$, $r>j$. Therefore

$$b_r = \sum_{j=r}^{p} \lambda_{jr} B_j \text{ , } r = q+1, \ldots, p \text{, involves only } B_r, B_{r+1}, \ldots, B_p .$$

Similarly, from equation (24)

$$B_j = \sum_{i=j}^{p} \mathcal{C}_{ij} b_j \text{ , } j = q+1, \ldots, p \text{, involves only } b_j, b_{j+1}, \ldots, b_p .$$

By equation (19)

$$c_{rn} = \sum_{j,m=1}^{p} \lambda_{jr} \lambda_{mn} C_{jm} \text{ , but}$$

$\lambda_{jr} = 0$, $r>j$ and $\lambda_{mn} = 0$, $n>m$. Therefore

$$c_{rn} = \sum_{j=r}^{p} \sum_{m=n}^{p} \lambda_{jr} \lambda_{mn} C_{jm} \text{ , } r, n = q+1, \ldots, p \text{, involves only those}$$

values of C_{jm} in which $j = r, r+1, \ldots, p$ and $m=n, n+1, \ldots, p$.

Similarly, from equation (20)

$$C_{jm} = \sum_{k=j}^{p} \sum_{n=m}^{p} c_{kn} \mathcal{C}_{kj} \mathcal{C}_{nm} \text{ , } \quad j, m = q+1, \ldots, p \text{, involves only}$$

those values of c_{kn} in which $k = j, j+1, \ldots, p$ and $n=m, m+1, \ldots, p$.

A similar relationship does not hold for the a_{ij} as now defined, however, where

$$\sum_{j=1}^{p} a_{ij}c_{jk} = \delta_{ik} \quad , \quad i,k = 1,2,\ldots, p \; .$$

For the reduced set of normal equations corresponding to x_{q+1}, $x_{q+2},\ldots,x_p$ let a^*_{ij} be defined such that

$$(25) \quad \sum_{r=q+1}^{p} a^*_{ir}c_{rs} = \delta_{is} \qquad i,s = q+1,\ldots,p. \quad \text{Also}$$

$$(26) \quad \sum_{j=q+1}^{p} A^*_{ij}C_{jk} = \delta_{ik} \qquad i,k = q+1,\ldots,p.$$

It has been shown that the same relations exist between b_r and B_r and between c_{rn} and C_{rn} for the subset $x_{q+1},\ldots, x_p$, as held in the original system of p independent variables. If the same thing can be proved in the case of a^*_{ij} and A^*_{ij} as compared with a_{ij} and A_{ij}, the proof will be completed, and formula (13) will hold with a^*_{ij} and A^*_{ij} substituted for a_{ij} and A_{ij}.

Substituting for c_{rs} in (25), we have

$$\sum_{r=q+1}^{p} \sum_{j=r}^{p} \sum_{m=s}^{p} a^*_{ir} \lambda_{jr} \lambda_{ms} C_{jm} = \delta_{is} \; .$$

Multiplying by μ_{st} and summing we obtain

$$\sum_{r,s=q+1}^{p} \sum_{j=r}^{p} \sum_{m=s}^{p} a^*_{ir} \lambda_{jr} \lambda_{ms} \mu_{st} C_{jm} = \sum_{s=q+1}^{p} \mu_{st} \delta_{is} \; ,$$

which from the previous relations gives

$$\sum_{r=q+1}^{p} \sum_{j=r}^{p} \sum_{m=s}^{p} a^*_{ir} \lambda_{jr} \delta_{mt} C_{jm} = \mathcal{C}_{it},$$

and summation over m, gives

$$\sum_{r=q+1}^{p} \sum_{j=r}^{p} a^*_{ir} \lambda_{jr} C_{jt} = \mathcal{C}_{it} \quad .$$

Then multiplying by A^*_{tn}, we have

$$\sum_{r=q+1}^{p} \sum_{j=r}^{p} \sum_{t=q-1}^{p} a^*_{ir} \lambda_{jr} C_{jt} A^*_{tn} = \sum_{t=q+1}^{p} \mathcal{C}_{it} A^*_{tn} \; .$$

Summing the left side over t gives

$$\sum_{r=q+1}^{p} \sum_{j=r}^{p} a^*_{ir} \lambda_{jr} \delta_{jn} \quad ,$$

and by summation over j we obtain

$$\sum_{r=q+1}^{p} a^*_{ir} \lambda_{nr} \quad .$$

Finally, multiplying by $\mathcal{C}_{vn}$ and summing we have

$$\sum_{r,n=q+1}^{p} a^*_{ir} \lambda_{nr} \mathcal{C}_{vn} = \sum_{t,n=q+1}^{p} \mathcal{C}_{it} A^*_{tn} \mathcal{C}_{vn} \quad .$$

Thus, we have

$$\sum_{r=q+1}^{p} a^*_{ir} \delta_{rv} = a^*_{iv},$$ and we have established the relationship

$$a^*_{iv} = \sum_{t,n=q+1}^{p} \mathcal{C}_{vn} \mathcal{C}_{it} A^*_{tn} \quad .$$

A comparison with equation (18) shows that this is exactly the same sort of relationship that exists between the a_{ij} and the A_{ij} in the original system which includes $x_1, x_2, \ldots, x_p$. It is now possible to verify equation (13) in its modified form and over the range x_{q+1} ,

$x_{q+2}, \ldots, x_p$.

$$\sum_{i,v=q+1}^{p} a^*_{iv} b_i b_v = \sum_{i,v,t,n=q+1}^{p} \sum_{j=1}^{p} \sum_{k=v}^{p} \mathcal{C}_{vn} \mathcal{C}_{it} A^*_{tn} \lambda_{ji} B_j \lambda_{kv} B_k .$$

Summation over i and v gives

$$= \sum_{t,n=q+1}^{p} \sum_{j=i}^{p} \sum_{k=v}^{p} \delta_{nk} \delta_{tj} A^*_{tn} B_j B_k ,$$

from which we obtain

$$\sum_{i,v=q+1}^{p} a^*_{iv} b_i b_v = \sum_{j,k=q+1}^{p} A^*_{jk} B_j B_k .$$

IV Conclusions

In conclusion, it has been proved that the reduction in sum of squares associated with a regression on the sub set of the independent variables $x_{q+1}, x_{q+2}, \ldots, x_p$ can be computed by the following four formulas:

(11) $S(y-Y')^2 - S(y-Y)^2 \equiv S(Y-Y')^2$

(12) $\sum_{i,j=q+1}^{p} a^*_{ij} b_i b_j$

(13) $\sum_{i=q+1}^{p} b_i g^*_i$

(14) $\sum_{i,j=q+1}^{p} c_{ij} g^*_i g^*_j$

Formula (11) represents the direct method which has ordinarily been used. The three alternative solutions discussed in this paper provide useful supplements to this method, in that they apply directly in some cases where the use of (11) would require considerable extra computation.

Bibliography

Nair, K. R. , "A Note on the Method of Fitting of Constants" Sankhya, Vol. 5 pt. 3, 1941

Yates, F. , "Orthogonal Functions and Tests of Significance in the Analysis of Variance, " Supp. Jour. Royal Statistical Society, Vol. V, 1938.