PEARSONIAN CORRELATION COEFFICIENTS ASSOCIATED WITH LEAST SQUARES THEORY

By Paul S. Dwyer

University of Michigan

1. Introduction and summary. It is well known that the zero-order correlation between the predicted value of a variable and the observed value of the variable is the multiple correlation. It is also well known that the zero-order correlation between the residuals for two different variables, when the prediction is from a common set of variables, is the partial correlation. These considerations naturally lead to a systematic investigation of all the zero-order correlations involving the various variables associated with least squares theory. Such an investigation is the purpose of this paper.

As a result of this study it appears that other zero-order correlations include the multiple alienation coefficient, the part correlation coefficient, and certain other coefficients which, as far as I am aware, have not been previously defined.

The paper first examines the case of a single predicted variable and then continues with the case in which two or more variables are predicted simultaneously. The paper includes (1) a theoretical development of the different coefficients and the relations between them, (2) the expression of the formulas in determinantal form, (3) a matrix presentation of the material, and (4) an outline of the calculational techniques—with illustrations.

It should be made clear at the start that this paper deals with populations (finite or infinite) and not with samples from those populations. The sampling distribution of each of the new correlation coefficients defined in this paper might well become the subject of a later investigation, but first we need to know what these correlation coefficients are.

2. The case of the single predicted variable. Notation, definitions, and basic properties. We suppose that a population consists of N individuals with values $X_{1j}, X_{2j}, \cdots, X_{kj}, Y_j$ for the variables $X_1, X_2, \cdots, X_k, Y$ and that Y is linearly predicted from the X_i by the formula

$$(1) \qquad E = Y - \alpha_0 - \alpha_1 X_1 - \alpha_2 X_2 - \cdots - \alpha_k X_k = Y - \underline{Y}$$

by least squares theory. For the purposes of this paper, we use a concise summation notation, ΣQ, in place of the more formal serial notation $\sum_{i=1}^{N} Q_i$ which is preferable to the frequency notation $\sum_{x=a}^{b} Q_x f_x$ and, in the continuous case, $\int_a^b Q_x f_x dx$ Moreover it is desirable that the scales of X and Y be chosen so as to facilitate the easy determination of the various formulas. If we let

$$y_j = \frac{Y_j - \bar{Y}}{\sqrt{N}\sigma_y}; \qquad x_{ij} = \frac{X_{ij} - \bar{X}_i}{\sqrt{N}\sigma_{x_i}} \tag{2}$$

we have $\Sigma x_i^2 = \Sigma y^2 = 1$ with the resulting correlating formula

$$\rho_{x_i y} = \frac{\Sigma x_i y}{\sqrt{(\Sigma x_i^2)(\Sigma y^2)}} = \Sigma x_i y \quad \text{and} \quad \rho_{x_i x_j} = \Sigma x_i x_j. \tag{3}$$

The transformations (2) when applied to (1) give

$$e = \frac{E}{\sqrt{N}\sigma_Y} = y - (\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k) = y - \underline{y} \tag{4}$$

where the β's are standard regression coefficients and e is defined to be $\frac{E}{\sqrt{N}\sigma_Y}$. It is to be noted that the values of x_i, y, e, and $\underline{y}$ are all dimensionless.

The values we wish to correlate are those of X_i, Y, E, $\underline{Y}$ of (1). The zero-order correlations involving these are the same as for x_i, y, e, $\underline{y}$ of (4).

3. Correlations with a single predicted variable. We wish to minimize Σe^2. Differentiating with respect to β_i and equating to zero we get

$$\Sigma e x_i = 0 \tag{5}$$

from which by multiplication by β_i and summation for i,

$$\Sigma e \underline{y} = 0. \tag{6}$$

It follows that

$$\Sigma e^2 = \Sigma e(y - \underline{y}) = \Sigma e y = \Sigma(y - \underline{y})y = \Sigma y^2 - \Sigma y \underline{y} = 1 - \Sigma y \underline{y} \tag{7}$$
$$= 1 - \Sigma(e + \underline{y})\underline{y} = 1 - \Sigma \underline{y}^2.$$

Using (4) and (7), we get

$$\Sigma e^2 = \frac{\Sigma E^2}{N\sigma_Y^2} = \frac{\sigma_E^2}{\sigma_Y^2} = 1 - \Sigma \underline{y}^2$$

so that

$$\Sigma \underline{y}^2 = \frac{\sigma_Y^2 - \sigma_E^2}{\sigma_Y^2}. \tag{8}$$

This is the conventional definition (from least squares theory) of the multiple correlation coefficient, so

$$\rho_{y:x_1 x_2 \cdots x_k}^2 = \rho_{y(x)}^2 = \Sigma \underline{y}^2 = \Sigma y \underline{y}. \tag{9}$$

Application of (9) to (7) gives

$$\Sigma e^2 = 1 - \rho_{y(x)}^2 = \kappa_{y(x)}^2 = \kappa_{y:x_1 x_2 \cdots x_k}^2 \tag{10}$$

where $\kappa_{y(x)}$ is the multiple alienation coefficient. We now have $\Sigma x_i^2 = 1$, $\Sigma y^2 = 1$, $\Sigma e^2 = \kappa_{y(x)}^2$, and $\Sigma \underline{y}^2 = \rho_{y(x)}^2$, so that we are able to present formulas involving $x_i, y, e, \underline{y}$. We first form the cross products

$$\Sigma xy = \rho_{xy}, \tag{11}$$

$$\Sigma xe = 0, \tag{12}$$

$$\Sigma x\underline{y} = \Sigma x(y + e) = \Sigma xy = \rho_{xy}, \tag{13}$$

$$\Sigma ye = \Sigma y(y - \underline{y}) = \Sigma y^2 - \Sigma y\underline{y} = 1 - \Sigma \underline{y}^2 = \kappa_{y(x)}^2, \tag{14}$$

$$\Sigma y\underline{y} = \Sigma \underline{y}^2 = \rho_{y(x)}^2, \tag{15}$$

$$\Sigma e\underline{y} = 0. \tag{16}$$

We then have

$$\rho_{x_i x_j} = \frac{\Sigma x_i x_j}{\sqrt{(\Sigma x_i^2)(\Sigma x_j^2)}} = \Sigma x_i x_j. \tag{17}$$

$$\rho_{x_i y} = \frac{\Sigma x_i y}{\sqrt{(\Sigma x_i^2)(\Sigma y^2)}} = \Sigma x_i y, \tag{18}$$

$$\rho_{xe} = \frac{\Sigma xe}{\sqrt{(\Sigma x^2)(\Sigma e^e)}} = 0, \tag{19}$$

$$\rho_{x\underline{y}} = \frac{\Sigma x\underline{y}}{\sqrt{(\Sigma x^2)(\Sigma \underline{y}^2)}} = \frac{\Sigma xy}{\rho_{y(x)}} = \frac{\rho_{xy}}{\rho_{y(x)}}. \tag{20}$$

It is interesting to note that this is unity in case $k = 1$ for then $\rho_{xy} = \rho_{y(x)}$. Otherwise the absolute value of $\rho_{x\underline{y}}$ is larger than that of ρ_{xy}. For this reason this coefficient might be called the *multiple augmented correlation coefficient.*

$$\rho_{ey} = \frac{\Sigma ey}{\sqrt{(\Sigma e^2)(\Sigma y^2)}} = \frac{\kappa_{y(x)}^2}{\kappa_{y(x)}} = \kappa_{y(x)}. \tag{21}$$

Thus the correlation between y and its residual is the multiple alienation coefficient.

$$\rho_{y\underline{y}} = \frac{\Sigma y\underline{y}}{\sqrt{(\Sigma y^2)(\Sigma \underline{y}^2)}} = \sqrt{\Sigma \underline{y}^2} = \rho_{y(x)}. \tag{22}$$

Thus, as is well known, the zero-order correlation between observed and predicted y is the multiple correlation.

$$\rho_{e\underline{y}} = \frac{\Sigma e\underline{y}}{\sqrt{(\Sigma e^2)(\Sigma \underline{y}^2)}} = 0. \tag{23}$$

4. Notation for the general case. We need to extend the notation and the definitions before examining explicit formulas for the more general case of two (or more) predicted variables. Suppose that Y_i and Y_j are the two variables predicted from the same X's. Then from (4) we write

$$(24)\qquad \begin{aligned} e_i &= \frac{E_i}{\sqrt{N}\sigma_{Y_i}} = y_i - \beta_{i1}x_1 - \beta_{i2}x_2 - \cdots - \beta_{ik}x_k = y_i - \underline{y}_i \\ e_j &= \frac{E_j}{\sqrt{N}\sigma_{Y_j}} = y_j - \beta_{j1}x_1 - \beta_{j2}x_2 - \cdots - \beta_{jk}x_k = y_j - \underline{y}_j . \end{aligned}$$

We then have the two sets of normal equations

$$(25)\qquad \Sigma e_i x = 0 \qquad \Sigma e_j x = 0$$

so that

$$(26)\qquad \begin{aligned} \Sigma e_i \underline{y}_i &= 0 \qquad \Sigma e_j \underline{y}_i = 0 \\ \Sigma e_i \underline{y}_j &= 0 \qquad \Sigma e_j \underline{y}_j = 0. \end{aligned}$$

It follows that

$$(27)\qquad \begin{aligned} \Sigma e_i e_j &= \Sigma e_i (y_j - \underline{y}_j) = \Sigma e_i y_j = \Sigma (y_i - \underline{y}_i) y_j = \Sigma y_i y_j - \Sigma \underline{y}_i y_j \\ &= \Sigma y_i y_j - \Sigma y_i \underline{y}_j = \Sigma y_i y_j - \Sigma \underline{y}_i \underline{y}_j = \rho_{ij} - \Sigma \underline{y}_i \underline{y}_j \end{aligned}$$

if we use the notation that $\rho_{ij} = \rho_{y_i y_j}$.

5. The correlations involving more than one predicted variable. In this case the y's, the e's and the $\underline{y}$'s (as well as the x's) can have more than one variable so that the correlation coefficients we need, in addition to those of section 3, are $\rho_{y_i y_j}$, $\rho_{e_i e_j}$, $\rho_{\underline{y}_i \underline{y}_j}$, $\rho_{y_i e_j}$, $\rho_{y_j e_i}$, $\rho_{y_i \underline{y}_j}$, $\rho_{\underline{y}_i y_j}$, $\rho_{e_i \underline{y}_j}$, and $\rho_{\underline{y}_i e_j}$. We need now only the summed products

$$(28)\qquad \begin{aligned} \Sigma y_i y_j &= \rho_{y_i y_j} = \rho_{ij}, \\ \Sigma e_i e_j &= \rho_{ij} - \Sigma \underline{y}_i \underline{y}_j \qquad \text{as given in (27)}, \end{aligned}$$

$$(29)\qquad \Sigma y_i e_j = \Sigma y_i (y_j - \underline{y}_j) = \Sigma y_i y_j - \Sigma y_i \underline{y}_j = \rho_{ij} - \Sigma \underline{y}_i \underline{y}_j,$$

$$(30)\qquad \Sigma y_i \underline{y}_j = \Sigma \underline{y}_i \underline{y}_j,$$

$$(31)\qquad \Sigma e_i \underline{y}_j = 0.$$

We have then

$$(32)\qquad \rho_{ij} = \frac{\Sigma y_i y_j}{\sqrt{(\Sigma y_i^2)(\Sigma y_j^2)}} = \Sigma y_i y_j ,$$

$$(33)\qquad \rho_{e_i e_j} = \frac{\Sigma e_i e_j}{\sqrt{(\Sigma e_i^2)(\Sigma e_j^2)}} = \frac{\rho_{ij} - \Sigma \underline{y}_i \underline{y}_j}{\kappa_{i(x)} \kappa_{j(x)}} .$$

This is the partial correlation coefficient.

$$(34)\qquad \rho_{\underline{y}_i \underline{y}_j} = \frac{\Sigma \underline{y}_i \underline{y}_j}{\sqrt{(\Sigma \underline{y}_i^2)(\Sigma \underline{y}_j^2)}} = \frac{\Sigma \underline{y}_i \underline{y}_j}{\rho_{i(x)} \rho_{j(x)}} .$$

This coefficient appears to be new. Since it is the correlation of predicted values, I suggest that it be called the *predictions correlation coefficient.*

$$\rho_{y_i e_j} = \frac{\Sigma y_i e_j}{\sqrt{(\Sigma y_i^2)(\Sigma e_j^2)}} = \frac{\rho_{ij} - \Sigma \underline{y}_i \underline{y}_j}{\kappa_{j(x)}}, \tag{35}$$

$$\rho_{e_i y_j} = \frac{\Sigma e_i y_j}{\sqrt{(\Sigma e_i^2)(\Sigma y_j^2)}} = \frac{\rho_{ij} - \Sigma \underline{y}_i \underline{y}_j}{\kappa_{i(x)}}. \tag{36}$$

The correlations given by (35) and (36) have been defined previously and are known as part correlation coefficients [1; 213,497].

$$\rho_{y_i \underline{y}_j} = \frac{\Sigma \underline{y}_i \underline{y}_j}{\sqrt{(\Sigma y_i^2)(\Sigma \underline{y}_j^2)}} = \frac{\Sigma \underline{y}_i \underline{y}_j}{\rho_{j(x)}}, \tag{37}$$

$$\rho_{\underline{y}_i y_j} = \frac{\Sigma \underline{y}_i \underline{y}_j}{\sqrt{(\Sigma \underline{y}_i^2)(\Sigma y_j^2)}} = \frac{\Sigma \underline{y}_i \underline{y}_j}{\rho_{i(x)}}. \tag{38}$$

The correlations of (37) and (38) appear to be new. Each is, in a sense, a generalization of the multiple correlation coefficient since it becomes the multiple correlation coefficient when $i = j$. I suggest that it might be called the *cross multiple correlation coefficient,* since it correlates the actual value of one variable with the predicted value of another.

$$\begin{aligned} \rho_{e_i \underline{y}_j} &= \frac{\Sigma e_i \underline{y}_j}{\sqrt{(\Sigma e_i^2)(\Sigma \underline{y}_j^2)}} = 0, \\ \rho_{\underline{y}_i e_j} &= \frac{\Sigma \underline{y}_i e_j}{\sqrt{(\Sigma \underline{y}_i^2)(\Sigma e_j^2)}} = 0. \end{aligned} \tag{39}$$

A summary of definitions and names of Pearsonian correlation coefficients associated with least squares theory is presented in Table I. No name is proposed when the coefficient is identically zero.

6. Relations between the correlations. Many relations exist between the correlations defined in earlier sections. Some of the more interesting of these are obtained by the elimination of $\Sigma \underline{y}_i \underline{y}_j$ from formulas involving this term. Thus from (34), (37), and (38) we get

$$\Sigma \underline{y}_i \underline{y}_j = \rho_{\underline{y}_i \underline{y}_j} \rho_{i(x)} \rho_{j(x)} = \rho_{y_i \underline{y}_j} \rho_{j(x)} = \rho_{\underline{y}_i y_j} \rho_{i(x)},$$

and from (33), (35), and (36) we get

$$\rho_{ij} - \Sigma \underline{y}_i \underline{y}_j = \rho_{e_i e_j} \kappa_{i(x)} \kappa_{j(x)} = \rho_{y_i e_j} \kappa_{j(x)} = \rho_{e_i y_j} \kappa_{i(x)}.$$

We then have

$$\left.\begin{aligned} &\rho_{ij} - \rho_{\underline{y}_i \underline{y}_j} \rho_{i(x)} \rho_{j(x)} \\ &\rho_{ij} - \rho_{y_i \underline{y}_j} \rho_{j(x)} \\ &\rho_{ij} - \rho_{\underline{y}_i y_j} \rho_{i(x)} \end{aligned}\right\} = \left\{\begin{aligned} &\rho_{e_i e_j} \kappa_{i(x)} \kappa_{j(x)} \\ &\rho_{y_i e_j} \kappa_{j(x)} \\ &\rho_{e_i y_j} \kappa_{i(x)} \end{aligned}\right. \tag{40}$$

where the six members may be equated in all possible ways.

Interesting and simple relations can also be obtained by formation of ratios. Thus

$$(41)\qquad \frac{\rho_{e_ie_j}}{\rho_{y_ie_j}} = \frac{1}{\kappa_{i(x)}}, \qquad \frac{\rho_{e_ie_j}}{\rho_{e_iy_j}} = \frac{1}{\kappa_{j(x)}} \qquad \text{so } \frac{\rho_{y_ie_j}}{\rho_{e_iy_j}} = \frac{\kappa_{i(x)}}{\kappa_{j(x)}}.$$

TABLE I

Definition	Name
Single predicted variable	
$\rho_{x_ix_j}$	Correlation coefficient of zero order
ρ_{xy}	Correlation coefficient of zero order
$\rho_{xe} = 0$	None
$\rho_{x\underline{y}} = \dfrac{\rho_{xy}}{\rho_{y\underline{x}}}$	*Multiple augmented correlation coefficient
$\rho_{ye} = \kappa_{y(x)}$	Multiple alienation coefficient
$\rho_{y\underline{y}} = \rho_{y(x)}$	Multiple correlation coefficient
$\rho_{e\underline{y}} = 0$	None
Two or more predicted variables	
$\rho_{y_iy_j}$	Correlation coefficient of zero order
$\rho_{e_ie_j}$	Partial correlation coefficient
$\rho_{\underline{y}_i\underline{y}_j}$	*Predictions correlation coefficient
$\rho_{y_ie_j}$	Part correlation coefficient
$\rho_{y_i\underline{y}_j}$	*Cross multiple correlation coefficient
$\rho_{e_i\underline{y}_j}$	None

* Proposed name

Similarly

$$(42)\qquad \frac{\rho_{\underline{y}_iy_j}}{\rho_{y_i\underline{y}_j}} = \frac{\rho_{i(x)}}{\rho_{j(x)}}.$$

The geometric mean of similar coefficients yields such expressions as

$$(43)\qquad \begin{aligned} \sqrt{\rho_{y_ie_j}\,\rho_{e_iy_j}} &\doteq \rho_{e_ie_j}\sqrt{\kappa_{i(x)}\,\kappa_{j(x)}} \\ \sqrt{\rho_{y_i\underline{y}_j}\,\rho_{\underline{y}_iy_j}} &= \rho_{y_iy_j}\sqrt{\rho_{i(x)}\,\rho_{j(x)}}. \end{aligned}$$

7. Determinantal formulas. The implicit normal equations (5) become when expanded

$$(44)\qquad \begin{aligned} \rho_{11}\beta_1 + \rho_{12}\beta_2 + \cdots + \rho_{1k}\beta_k &= \rho_{1y} \\ \rho_{21}\beta_1 + \rho_{22}\beta_2 + \cdots + \rho_{2k}\beta_k &= \rho_{2y} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \rho_{k1}\beta_1 + \rho_{k2}\beta_2 + \cdots - \rho_{kk}\beta_k &= \rho_{ky} \end{aligned}$$

while $\Sigma yy = \Sigma y^2 = \rho^2_{y(x)}$ becomes

$$\rho_{y1}\beta_1 + \rho_{y2}\beta_2 + \cdots + \rho_{yk}\beta_k = \rho^2_{y(x)}. \tag{45}$$

Let Δ be the determinant of the matrix of the solution of the k x's and y. Let Δ' be the corresponding determinant with ρ_{yy} replaced by $\rho^2_{y(x)}$. Let Δ_{yy} be the determinant of the correlation matrix of the k x's. Then $\rho^2_{y(x)} = \Sigma y^2 = \Sigma yy$ can be expressed as a function of Δ and Δ_{yy}. If (44) and (45) are to hold simultaneously, then $\Delta' = 0$. Expanding Δ' in terms of the bottom row, we get

$$\Delta' = 0 = \rho^2_{y(x)}\Delta_{yy} + \text{“terms”}. \tag{46}$$

Similarly

$$\Delta = \rho_{yy}\Delta_{yy} + \text{“terms”} \tag{47}$$

where the "terms" of (46) and (47) are identical. It follows by subtraction that $\Delta = (1 - \rho^2_{y(x)})\Delta_{yy}$ and hence that

$$\Sigma yy = \Sigma y^2 = \rho^2_{y(x)} = 1 - \frac{\Delta}{\Delta_{yy}}. \tag{48}$$

Then

$$\Sigma e^2 = \Sigma ey = \kappa^2_{y(x)} = 1 - \Sigma y^2 = 1 - \left(1 - \frac{\Delta}{\Delta_{yy}}\right) = \frac{\Delta}{\Delta_{yy}}. \tag{49}$$

Correlation formulas of section 3 then appear as

$$\rho_{xy} = \frac{\rho_{xy}}{\sqrt{1 - \frac{\Delta}{\Delta_{yy}}}}, \tag{50}$$

$$\rho_{ey} = \sqrt{\frac{\Delta}{\Delta_{yy}}}, \tag{51}$$

$$\rho_{yy} = \sqrt{1 - \frac{\Delta}{\Delta_{yy}}}. \tag{52}$$

In a similar way the normal equations (25) become two sets of normal equations. The first set is like (44) with β_s replaced by β_{y_is} and ρ_{ys} replaced by ρ_{y_is}. The second set is similar with i replaced by j. It is desired to find

$$\Sigma y_i y_j = \Sigma y_i y_j = \rho_{y_j1}\beta_1 + \rho_{y_j2}\beta_2 + \cdots + \rho_{y_jk}\beta_k. \tag{53}$$

Now using (53) with (51) as applied to y_i and using the technique of the first part of this section, we get

$$\Delta_{y_iy_j} = \rho_{y_iy_j}\Delta_{y_iy_i\cdot y_jy_j} + \text{“terms”}, \tag{54}$$

$$0 = \Sigma y_i y_j \Delta_{y_iy_i\cdot y_jy_j} + \text{“terms”}, \tag{55}$$

where Δ is the determinant of the matrix of the correlations of the k x's, y_i and

y_j; $\Delta_{y_i y_j}$ is the determinant obtained by deleting the column involving correlations of y_i and the row involving correlations of y_i; $\Delta_{y_i y_i \cdot y_j y_j}$ is the determinant of the matrix of the k x's; and the "terms" in (54) and (55) are identical. It follows that

(56)
$$\Sigma \underline{y}_i \underline{y}_j = \rho_{ij} - \frac{\Delta_{y_j y_i}}{\Delta_{y_i y_i \cdot y_j y_j}}$$

and thence

(57)
$$\rho_{ij} - \Sigma \underline{y}_i \underline{y}_j = \frac{\Delta_{y_i y_j}}{\Delta_{y_i y_i \cdot y_j y_j}}.$$

The formulas of section (5) then appear in determinant form as follows

(58)
$$\rho_{e_i e_j} = \frac{\dfrac{\Delta_{ij}}{\Delta_{ii \cdot jj}}}{\sqrt{\left(\dfrac{\Delta_{ii}}{\Delta_{ii \cdot jj}}\right)\left(\dfrac{\Delta_{jj}}{\Delta_{jj \cdot ii}}\right)}} = \frac{\Delta_{ij}}{\sqrt{\Delta_{ii} \Delta_{jj}}}$$

as is well known.

(59)
$$\rho_{\underline{y}_i \underline{y}_j} = \frac{\rho_{ij} - \dfrac{\Delta_{ij}}{\Delta_{ii \cdot jj}}}{\sqrt{\left(1 - \dfrac{\Delta_{ii}}{\Delta_{ii \cdot jj}}\right)\left(1 - \dfrac{\Delta_{jj}}{\Delta_{jj \cdot ii}}\right)}}.$$

(60)
$$\rho_{y_i e_j} = \frac{\dfrac{\Delta_{ij}}{\Delta_{ii \cdot jj}}}{\sqrt{\dfrac{\Delta_{ij}}{\Delta_{ii \cdot jj}}}}.$$

(61)
$$\rho_{y_i \underline{y}_j} = \frac{\rho_{ij} - \dfrac{\Delta_{ij}}{\Delta_{ii \cdot jj}}}{\sqrt{1 - \dfrac{\Delta_{ii}}{\Delta_{ii \cdot jj}}}}.$$

Formulas for $\rho_{e_i y_j}$ and $\rho_{\underline{y}_i y_j}$ are similar to (60) and (61).

Modern methods of calculating determinants (2), (3), (4), (5) are advised if calculations are to be made from those formulas.

8. Matrix formulas. A matrix presentation is very useful in exhibiting the general features of this theory and in developing compact and easy methods of calculation with finite populations. The matrix presentation here is similar to that given by the author in a previous article [6].

Let the normal equations (24) be represented by the matrix equation.

(62)
$$E = Y - XB = Y - \underline{Y}.$$

Then the sets of normal equations become

$$X'E = 0 \quad \text{or} \quad X'(Y - XB) = 0$$

so that

$$X'XB = X'Y. \tag{63}$$

Now since $XB = \underline{Y}$, (63) can be written as $X'\underline{Y} = X'Y$ and it can be shown that

$$\underline{Y}'Y = Y'\underline{Y} = \underline{Y}'\underline{Y}. \tag{64}$$

But under the assumptions of section 2, $X'X$ is the matrix of the intercorrelations of the X's, $X'Y$ is the matrix of the intercorrelations of the x's and y's and $Y'Y$ is the matrix of the intercorrelations of the y's. Hence (63) can be written

$$R_{xx}B = R_{xy} \tag{65}$$

so that

$$B = R_{xx}^{-1}R_{xy}. \tag{66}$$

If Y is composed of a single variable, B is a single column matrix (vector) but if Y is composed of m variables, B is an m column matrix. It follows at once that

$$\underline{Y}'Y = \underline{Y}'\underline{Y} = B'X'XB = B'R_{xy}B = R_{xy}'R_{xx}^{-1}R_{xx}R_{xx}^{-1}R_{xy} = R_{xy}'R_{xx}^{-1}R_{xy} \tag{67}$$

and that

$$\begin{aligned} E'E = (Y - XB)'E = Y'E = Y'(Y - XB) &= Y'Y - Y'\underline{Y} \\ &= Y'Y - \underline{Y}'\underline{Y} = R_{yy} - R_{xy}'R_{xx}^{-1}R_{xy}. \end{aligned} \tag{68}$$

It thus appears that the matrix (67) has diagonal terms $\Sigma \underline{y}^2 = \Sigma y\underline{y}$ which are the squares of the multiple correlation coefficients, and that the non-diagonal terms are $\Sigma \underline{y}_i\underline{y}_j = \Sigma y_i\underline{y}_j$. Similarly the matrix (68) has diagonal terms $\Sigma e^2 = \kappa_{y(x)}^2$ and non-diagonal terms $\Sigma e_ie_j = \Sigma e_iy_j$. It follows that all the correlation coefficients defined above may be calculated from the matrices R_{xx}, R_{xy}, R_{yy}, $\underline{Y}'\underline{Y}$, and $E'E$. The matrix (67) might be called the *multiple correlation matrix* and the matrix (68) the *multiple alienation matrix*.

Conventional results are expressed in terms of the correlation matrices R_{xx}, R_{xy}, and R_{yy}. All the correlation coefficients defined in this paper may be expressed in terms of these matrices and the multiple correlation and alienation matrices.

9. Calculational method of determining the multiple correlation and multiple alienation matrices. Various methods might be used in calculating the multiple correlation and alienation matrices from the correlation matrices. One method utilizes the square root method of solving simultaneous equations, which has recently been presented in a number of places, [7] [8] together with a device which is similar to that used by Aitken [9] in eliminating the back solution. This method solves the equation (65) by forming the auxiliary

$$S_{xx}B = S_{xx}R_{xx}^{-1}R_{xy} \tag{69}$$

where S_{xx} is a triangular matrix such that

$$R_{xx} - S'_{xx}S_{xx} = 0. \tag{70}$$

TABLE II

General		Illustration					
	R_{yy}					1.000	.495
						—	1.000
		1.000	.652	.554	.615	.313	.650
		—	1.000	.747	.693	.280	.803
R_{xx}	R_{xy}	—	—	1.000	.774	.182	.804
		—	—	—	1.000	.166	.812
		1.000	.652	.554	.615	.313	.650
			.758	.509	.385	.100	.500
S_{xx}	$S_{xx}R_{xx}^{-1}R_{xy}$			.659	.360	.064	.287
					.586	.072	.199
						.117	.221
	$\underline{Y}'\underline{Y}$					—	.794
						.883	.274
	$E'E$					—	.206

The right hand side of (69), when premultiplied by its transpose yields

$$(S_{xx}R_{xx}^{-1}R_{xy})'(S_{xx}R_{xx}^{-1}R_{xy}) = R'_{xy}R_{xx}^{-1}S'_{xx}S_{xx}R_{xx}^{-1}R_{xy} = R'_{xy}R_{xx}^{-1}R_{xy} = \underline{Y}'\underline{Y}. \tag{71}$$

Speaking less technically it is only necessary to multiply the columns of $S_{xx}R_{xx}^{-1}R_{xy}$ to get $\underline{Y}'\underline{Y}$.

A first illustration utilizes the correlations of the Carver anthropometric data [10] for 1000 University of Michigan freshmen. This group may be regarded as constituting a population, or it may be regarded as a random sample of a larger population. For present purposes we regard it as a population. Height (Y_1) and weight (Y_2) are estimated from shoulder girth (X_1) chest girth (X_2), waist girth (X_3), and right thigh girth (X_4). The calculation of $\underline{Y}'\underline{Y}$ and $E'E$ from the correlation matrices follow.

As a second illustration I use the correlation between the parts of two forms of the Thorndike Intelligence Examination which Lorge has used in illustration canonical correlation technique [11, 69–74]. The X's are the scores on the three parts of Form A and the Y's are the scores on the three parts of Form B. In this case we designate the results by r's and k's (rather than ρ's and κ's) since the calculation is considered to be for a sample. The calculation of the sample multiple correlation and multiple alienation matrices is presented in Table III.

TABLE III

	Form A			Form B			
	x_1	x_2	x_3	y_1	y_2	y_3	
				1.0000	.8235	.7912	
				—	1.0000	.8315	R_{yy}
				—	—	1.0000	
	1.0000	.7830	.7852	.8986	.7841	.8217	
R_{xx}	—	1.0000	.8393	.7961	.8543	.8254	R_{xy}
	—	—	1.0000	.7683	.8226	.8588	
	1.0000	.7830	.7852	.8986	.7841	.8217	
S_{xx}		.6220	.3609	.1487	.3864	.2926	$S_{xx}R_{xx}^{-1}R_{xy}$
			.5032	.0180	.1341	.2146	
				.8299	.7645	.7858	
				—	.7821	.7861	$Y'Y$
				—	—	.8069	
				.1701	.0590	.0054	
				—	.2179	.0454	$E'E$
				—	—	.1991	

10. The numerical values of the coefficients. The diagonal entries of the multiple correlation matrix give the values of $\Sigma y_i^2 = \Sigma y_i y_i = \rho^2_{y(x)}$ while the non-diagonal values are $\Sigma y_i y_j = \Sigma y_i y_j$. The diagonal entries of the multiple alienation matrix are $\Sigma e_i^2 = \Sigma e_i y_i = \kappa^2_{y(x)}$ while the non-diagonal entries are $\Sigma e_i e_j = \Sigma e_i y_j = \Sigma y_i e_j$. We are then able to write out any of the correlations easily. Thus from Table II

$$\rho_{1(x)} = \sqrt{\Sigma y_1^2} = \sqrt{.117} = .342,$$

$$\rho_{2(x)} = \sqrt{\Sigma y_2^2} = \sqrt{.794} = .891,$$

$$\kappa_{1(x)} = \sqrt{\Sigma e_1^2} = \sqrt{.883} = .940,$$

$$\kappa_{2(x)} = \sqrt{\Sigma e_2^2} = \sqrt{.206} = .454,$$

$$\rho_{12(x)} = \frac{\Sigma e_1 e_2}{\sqrt{(\Sigma e_1^2)(\Sigma e_2^2)}} = \frac{.274}{\sqrt{(.883)(.206)}} = .643,$$

$$\rho_{\underline{y}_1 \underline{y}_2} = \frac{\Sigma \underline{y}_1 \underline{y}_2}{\sqrt{(\Sigma \underline{y}_1^2)(\Sigma \underline{y}_2^2)}} = \frac{.221}{\sqrt{(.117)(.794)}} = .724,$$

$$\rho_{y_1 e_2} = \frac{\Sigma e_1 e_2}{\sqrt{\Sigma e_2^2}} = \frac{.274}{\sqrt{.206}} = .604,$$

$$\rho_{y_2 e_1} = \frac{\Sigma e_1 e_2}{\sqrt{\Sigma e_1^2}} = \frac{.274}{\sqrt{.883}} = .291,$$

$$\rho_{y_1 y_2} = \frac{\Sigma \underline{y}_1 \underline{y}_2}{\sqrt{\Sigma \underline{y}_2^2}} = \frac{.221}{\sqrt{.794}} = .248,$$

$$\rho_{y_1 y_2} = \frac{\Sigma \underline{y}_1 \underline{y}_2}{\sqrt{\Sigma \underline{y}_1^2}} = \frac{.221}{\sqrt{.117}} = .646.$$

TABLE IVa

<table>
<tr><th colspan="3">General</th><th colspan="3">Illustration</th></tr>
<tr><td>$\rho_{1(x)}$
$\Sigma \underline{y}_1^2$</td><td>$r_{y_1y_2}$
$r_{y_1y_2} \mid r_{y_1y_2}$
$\Sigma \underline{y}_1\underline{y}_2$</td><td>$r_{y_1y_3}$
$r_{y_1y_3} \mid r_{y_1y_3}$
$\Sigma \underline{y}_1\underline{y}_3$</td><td>.9110
.8299</td><td>.9489
.8392 | .8644
.7645</td><td>.9603
.8626 | .8747
.7858</td></tr>
<tr><td></td><td>$r_{2(x)}$
$\Sigma \underline{y}_2^2$</td><td>$r_{y_2y_3}$
$r_{y_2y_3} \mid r_{y_2y_3}$
$\Sigma \underline{y}_2\underline{y}_3$</td><td></td><td>.8844
.7821</td><td>.9917
.8889 | .8751
.7861</td></tr>
<tr><td></td><td></td><td>$r_{3(x)}$
$\Sigma \underline{y}_3^2$</td><td></td><td></td><td>.8983
.8069</td></tr>
</table>

TABLE IVb

<table>
<tr><th colspan="3">General</th><th colspan="3">Illustration</th></tr>
<tr><td>$k_{1(x)}$
Σe_1^2</td><td>$r_{e_1e_2}$
$r_{e_1y_2} \mid r_{y_1e_2}$
Σe_1e_2</td><td>$r_{e_1e_3}$
$r_{e_1y_3} \mid r_{y_1e_3'}$
Σe_1e_3</td><td>.4124
.1701</td><td>.3066
.1431 | .1264
.0590</td><td>.0298
.0131 | .0123
.0054</td></tr>
<tr><td></td><td>$k_{2(x)}$
Σe_2^2</td><td>$r_{e_2e_3}$
$r_{e_2y_3} \mid r_{y_2e_3}$
Σe_2e_3</td><td></td><td>.4668
.2179</td><td>.2214
.0973 | .1033
.0454</td></tr>
<tr><td></td><td></td><td>$k_{3(x)}$
Σe_3^2</td><td></td><td></td><td>.4394
.1931</td></tr>
</table>

It is possible to utilize a scheme of successive division if all these correlations are desired when there are more than two predicted variables. By divisions we compute in turn $\rho_{i(x)}$, $\rho_{\underline{y}_i y_j}$, $\rho_{y_i \underline{y}_j}$ and $\rho_{\underline{y}_i \underline{y}_j}$ from the multiple correlation matrix and $\kappa_{i(x)}$ $\rho_{e_i y_j}$, $\rho_{y_i e_j}$, $\rho_{e_i e_j}$ from the multiple alienation matrix for each i, j. The computational scheme is illustrated in Table IV where the correlations used are the sample correlations of Table III. The calculations from the multiple correlation matrix are presented in Table IVa and those from the multiple alienation matrix in Table IVb.

In Table IVa the multiple correlation matrix is first entered on the third of each three lines. The square root of each diagonal term is then extracted to give the multiple correlation coefficients. The value of $r_{i(x)}$ is then locked in the machine as a divisor and it is divided, in turn, into $\Sigma \underline{y}_1 y_2$, $\Sigma \underline{y}_1 y_3$ to get $r_{\underline{y}_1 y_2}$ and $r_{\underline{y}_1 y_3}$. Then $r_{2(x)}$ is used as a divisor by division into $r_{\underline{y}_1 y_2}$ to get $r_{\underline{y}_1 \underline{y}_2}$, into $\Sigma y_1 \underline{y}_2$ to get $r_{y_1 \underline{y}_2}$, and into $\Sigma \underline{y}_2 y_3$ to get $r_{\underline{y}_2 y_3}$. Finally $r_{3(x)}$ is divided into $r_{\underline{y}_1 y_3}$ to get $r_{\underline{y}_1 \underline{y}_3}$, into $\Sigma y_1 \underline{y}_3$ to get $r_{y_1 \underline{y}_3}$, into $r_{\underline{y}_2 y_3}$ to get $r_{\underline{y}_2 \underline{y}_3}$ and into $\Sigma y_2 \underline{y}_3$ to get $r_{y_2 \underline{y}_3}$. A check on these divisions can be made, if desired, by dividing $r_{y_1 \underline{y}_2}$ by $r_{1(x)}$ to get $r_{\underline{y}_1 \underline{y}_2}$, $r_{y_1 \underline{y}_3}$ by $r_{1(x)}$ to get $r_{\underline{y}_1 \underline{y}_3}$, and $r_{y_2 \underline{y}_3}$ by $r_{2(x)}$ to get $r_{\underline{y}_2 \underline{y}_3}$.

Table IVb is treated in a similar manner.

This technique is immediately applicable to the case of many predicted variables.

REFERENCES

[1] M. EZEKIEL, *Methods of Correlation Analysis*, Second Edition, Wiley, 1942.

[2] A. C. AITKEN, "On the evaluation of determinants, the formation of their adjugates, etc.," *Edinb. Math. Soc. Proc.*, Series 2, Vol. 3(1933), pp. 207–219.

[3] P. S. DWYER, "The evaluation of determinants," *Psychometrika*, Vol. 6(1941), pp. 191–204.

[4] A. C. AITKEN, *Determinants and Matrices*, Second Edition, Oliver and Boyd, Edinburgh, 1942.

[5] F. V. WAUGH AND P. S. DWYER, "Compact computation of the inverse of a matrix," *Annals of Math. Stat.*, Vol. 16(1945), pp. 359–371.

[6] P. S. DWYER, "A matrix presentation of least squares and correlation theory, etc.," *Annals of Math. Stat.*, Vol. 15(1944), pp. 82–89.

[7] P. S. DWYER, "The square root method and its use in correlation and regression," *Am. Stat. Assn. Jour.*, Vol. 40(1945), pp. 493–503.

[8] D. B. DUNCAN AND J. F. KENNEY, *On the Solution of Normal Equations and Related Topics*, Edwards Brothers, Ann Arbor, 1946.

[9] A. C. AITKEN, "The evaluation of a certain triple product matrix," *Roy. Soc. of Edinb. Proc.*, Vol. 57(1937), pp. 172–181.

[10] H. C. CARVER, *Anthropometric Data*, Edwards Brothers, Ann Arbor, 1941.

[11] IRVING LORGE, "The computation of Hotelling canonical correlations," *Proceedings of Educational Research Forum*, Endicott, N. Y., Aug. 26–31, 1940, pp. 68–74.