A CONCISE ANALYSIS OF CERTAIN ALGEBRAIC FORMS

By Franklin E. Satterthwaite

State University of Iowa, Iowa City, Iowa

Many of the statistics in common use are functions of homogeneous algebraic forms in the items of the sample. Among such statistics are the mean, a linear form; the variance, a quadratic form; and the product moment, a bilinear form. With the extension of the science, the mathematical statistician is faced with the study of more complex statistics and the associated algebraic forms and matrices. The purpose of this paper is to set forth concise and efficient notations and methods which may be used in such analysis.

We shall borrow the essential features of our notation from differential geometry and tensor analysis. The Kroneker delta is defined as,

$$\delta_i^j = 1, \qquad i = j;$$
$$= 0, \qquad i \neq j.$$

The summation convention provides that summation be performed with respect to any index appearing twice in the same term. Thus,

$$x_i y^i = x_1 y^1 + x_2 y^2 + \cdots .$$

To extend the use of the summation convention, we shall frequently place indices on the numeral, 1. Thus,

$$1^i x_i = 1^1 x_1 + 1^2 x_2 + \cdots = x_1 + x_2 + \cdots .$$

Symmetry in the calculations is more striking if the pair of summation indices appears, one as a superscript, the other as a subscript. Therefore we allow the shifting of an index from the one position to the other at will. Thus,

$$x_i \equiv x^i.$$

Where no confusion will arise, indices may be placed outside of parentheses.

$$\left(\delta\delta - \frac{1}{a}\frac{1}{b}\right)_{ik}^{jl} = \delta_i^j \delta_k^l - \left(\frac{1_i^j}{a}\right)\left(\frac{1_k^l}{b}\right).$$

The standard notations for averages will be used.

$$\bar{x}_{i\cdot} = \left(\frac{1}{b}\right)^j x_{ij} = \left(\frac{1}{b}\right)\Sigma_j x_{ij} \tag{1}$$

$$\bar{x} = \left(\frac{1}{N}\right)\Sigma x_{ijk\ldots} . \tag{2}$$

Unless otherwise indicated, the symbol, Σ, will always stand for summation over *all* unrepeated indices including any already averaged under conventions (1) and (2). Thus,

$$\Sigma \bar{x}^2 = N\bar{x}^2.$$

The following simple formulas are fundamental to the arithmetic of this notation. They are obvious upon the expansion of the summations. Each index varies from 1 to a. These formulas are

$$\delta_i^j x_j = x_i,$$

$$\delta_i^j \delta_j^k = \delta_i^k,$$

$$\delta_i^j 1_j^k = 1_i^k,$$

$$1_i^j 1_j^k = a1_i^k,$$

$$\delta_i^i = a,$$

$$x_i - \bar{x}_. = \delta_i^j x_j - \left(\frac{1}{a}\right)_i^j x_j,$$

$$= \left(\delta - \frac{1}{a}\right)_i^j x_j.$$

The symbols of this notation obey the associative, commutative, and the distributive laws of simple arithmetic so that the operations of summation, multiplication, and squaring are very easy. Thus for the product of two linear forms we have

$$\bar{x}\bar{y} = \left(\frac{1}{a}\right)^i x_i \left(\frac{1}{b}\right)^j x_j = \left(\frac{1}{ab}\right)_j^i x_i x^j.$$

The sum of squares is obtained by the simple repetition of the form,

$$\text{(3)} \quad \begin{aligned} \Sigma x_i^2 &= \Sigma(\delta_i^j x_j)^2 = (\delta_i^j x_j)(\delta_i^j x_j), \\ &= (\delta_i^j x_j)(\delta_k^i x^k) = \delta_k^j x_j x^k. \end{aligned}$$

Two other sums of squares occur so frequently that they should be particularly noted:

$$\text{(4)} \quad \begin{aligned} \Sigma \bar{x}^2 &= \Sigma\left[\left(\frac{1}{a}\right)_i^j x_j\right]^2 = \left(\frac{1}{a}\right)_i^j x_j \left(\frac{1}{a}\right)_k^i x^k, \\ &= \left(\frac{1}{a}\frac{1}{a}\right)_{ik}^{ji} x_j x^k = \left(\frac{1}{a}\right)_k^j x_j x^k. \end{aligned}$$

$$
\begin{aligned}
\Sigma(x_i - \bar{x})^2 &= \Sigma\left[\left(\delta - \frac{1}{a}\right)_i^j x_j\right]^2, \\
&= \left(\delta - \frac{1}{a}\right)_i^j \left(\delta - \frac{1}{a}\right)_k^i x_j x^k, \\
&= \left(\delta\delta - \frac{1}{a}\delta - \delta\frac{1}{a} + \frac{1}{a}\frac{1}{a}\right)_{ik}^{ji} x_j x^k, \qquad (5) \\
&= \left(\delta - \frac{1}{a} - \frac{1}{a} + \frac{1}{a}\right)_k^j x_j x^k, \\
&= \left(\delta - \frac{1}{a}\right)_k^j x_j x^k.
\end{aligned}
$$

The striking similarity in the coefficients of the second and final expressions for the summations in (3), (4), and (5) should not be overlooked.

Where we have multiple classification of the variables, we may operate on each index separately. For example, in a four-way analysis of variance we may have the quadratic form,

$$
\begin{aligned}
Q &= \Sigma\{\bar{x}_{ijk\cdot} - \bar{x}_{ij\cdot\cdot} - \bar{x}_{i\cdot k\cdot} + \bar{x}_{i\cdots}\}^2, \\
&= \Sigma\left\{\left[\delta\left(\delta\delta - \delta\frac{1}{c} - \frac{1}{b}\delta + \frac{1}{b}\frac{1}{c}\right)\frac{1}{d}\right]_{ijkl}^{mnop} x_{mnop}\right\}^2, \\
&= \Sigma\left\{\left[\delta\left(\delta - \frac{1}{b}\right)\left(\delta - \frac{1}{c}\right)\frac{1}{d}\right]_{ijkl}^{mnop} x_{mnop}\right\}^2, \\
&= \delta_q^m \left(\delta - \frac{1}{b}\right)_r^n \left(\delta - \frac{1}{c}\right)_s^o \left(\frac{1}{d}\right)_t^p x_{mnop} x^{qrst}.
\end{aligned}
$$

The rank is one of the important properties of a quadratic form or matrix. An experienced mathematician usually has a rule of thumb for determining the ranks of those quadratic forms occurring in statistical analysis. In order to formulate such rules of thumb into a simple and rigorous algebra, the author here defines a type of matrix multiplication which he calls "uncontracted matrix multiplication" and which he represents by the symbol, $\odot$.

Let $A = \| \alpha_i^j \|$ and $B = \| \beta_i^j \|$ be two matrices of any finite orders and with ranks R_A and R_B. We define the uncontracted product, $A \odot B$, as follows:

$$
\begin{aligned}
C &= A \odot B \\
&= \| \alpha_i^j \| \odot B \\
&= \| \alpha_i^j B \| \\
&= \begin{Vmatrix} \alpha_1^1 B & \alpha_1^2 B & \cdots \\ \alpha_2^1 B & \alpha_2^2 B & \cdots \\ \cdots & \cdots & \cdots \end{Vmatrix},
\end{aligned}
$$

where

$$\alpha_i^j L = \begin{Vmatrix} \alpha_i^j \beta_1^1 & \alpha_i^j \beta_1^2 & \cdots \\ \alpha_i^j \beta_2^1 & \alpha_i^j \beta_2^2 & \cdots \\ \cdots & \cdots & \cdots \end{Vmatrix}.$$

Thus the elements of C are

$$\gamma_{ik}^{jl} \equiv \alpha_i^j \beta_k^l .$$

We therefore see that whenever we have a matrix whose elements can be factored in the above manner, then the matrix can be expressed as the uncontracted product of simple matrices. Thus,

if $$\| \gamma_{ij\cdots}^{mn\cdots} \| \equiv \| (\alpha_i^m \beta_j^n \cdots) \|$$

then $$\| \gamma_{ij\cdots}^{mn\cdots} \| = \| \alpha_i^m \| \odot \| \beta_j^n \| \odot \cdots .$$

We shall now prove that the rank of the uncontracted product, $C = A \odot B$, of two matrices is equal to the product of the ranks. This follows because for the matrix, A, there always exists a set of elementary transformations defined by the equations,

$$T_A: \qquad {}_A\delta_i^j = \begin{pmatrix} j \\ s \end{pmatrix} \begin{pmatrix} r \\ i \end{pmatrix} \hat{\theta}_n^s \theta_r^m \alpha_m^n , \qquad \hat{\theta}_i^j ,\ \theta_i^j \neq 0,\ i = j,$$

where the θ_i^j's, $i = j$, are coefficients providing for the multiplication of the elements of a row by a constant not zero; the θ_i^j's, $i \neq j$, are coefficients providing for the addition to the elements of a row a linear function of the corresponding elements of the other rows; the $\hat{\theta}$'s are similar coefficients referring to columns; the symbol $\begin{pmatrix} j \\ i \end{pmatrix}$ is an operator indicating the interchange of the ith and jth rows (columns); and the ${}_A\delta$'s have the values,

$$\begin{aligned} {}_A\delta_i^j &= 1, & i = j \leq R_A , \\ &= 0, & \text{otherwise.} \end{aligned}$$

This set of transformations reduces A to a diagonal matrix with R_A non-zero elements. A similar set of transformations,

$$T_B: \qquad {}_B\delta_k^l = \begin{pmatrix} l \\ s \end{pmatrix} \begin{pmatrix} r \\ k \end{pmatrix} \hat{\varphi}_n^s \varphi_r^m \beta_m^n ,$$

exists for the matrix B. We next define two sets of transformations by the equations,

$$T_A': \qquad ({}_A\delta_i^j\ \beta_k^l) = \begin{pmatrix} jl \\ sl \end{pmatrix} \begin{pmatrix} rk \\ ik \end{pmatrix} \hat{\theta}_n^s \theta_r^m (\alpha_m^n\ \beta_k^l),$$

$$T_B': \qquad ({}_A\delta_i^j\ {}_B\delta_k^l) = \begin{pmatrix} jl \\ js \end{pmatrix} \begin{pmatrix} ir \\ ik \end{pmatrix} \hat{\varphi}_n^s \varphi_r^m ({}_A\delta_i^j\ \beta_m^n),$$

which are also elementary because of their relationship to T_A and T_B. Now if we subject the matrix, $C = \| (\alpha_i^j \beta_k^l) \|$ to the transformations T'_A followed by the transformations T'_B, it will be reduced to the diagonal form $C = \| ({}_A\delta_i^j{}_B\delta_k^l) \|$ with exactly $R_A R_B$ non-zero elements. Therefore, since the rank of a matrix is invariant under elementary transformations, the rank of $C = A \odot B$ must be $R_A R_B$.

We shall now determine the ranks of several matrices which occur frequently in statistics:

$$A_1 = \| 1_i \| = \| 1, 1, 1, \cdots \|, \quad R_1 = 1.$$

$$A_2 = \| 1_i^j \| = \| 1_i \cdot 1^j \| = \| 1_i \| \odot \| 1^j \|,$$

$$R_2 = 1 \cdot 1 = 1.$$

$$A_3 = \| \delta_i^j \|, \quad R_3 = a.$$

$$A_4 = \left\| \left(\delta - \frac{1}{a}\right)_i^j \right\|, \quad R_4 = a - 1.$$

The proof that $R_4 = a - 1$ involves two steps. First summing the rows of A_4 we have,

$$1^i \alpha_i^j = 1^i \delta_i^j - 1^i \left(\frac{1}{a}\right)_i^j = 1^j - \left(\frac{a}{a}\right)^j = 0$$

so that $R_4 \leq a - 1$. Second if we subtract the elements of the first row from the corresponding elements of each of the other rows we obtain,

$$A_4 \sim \left\| \begin{array}{c|c} 1 - \frac{1}{a} & \frac{1}{a} \\ \hline -1 & \delta_i^j \end{array} \right\| \begin{array}{l} i = 1 \\ i \neq 1. \end{array}$$

Since the $(a - 1)$st order determinant in the lower right-hand corner is not equal to zero, $R_4 \geq a - 1$.

Applying our theorem on uncontracted products, the ranks of complicated matrices can often be determined by inspection. Thus:

$$A_5 = \left\| \delta_i^j \left(\delta - \frac{1}{b}\right)_k^l \right\| = \| \delta_i^j \| \odot \left\| \left(\delta - \frac{1}{b}\right)_k^l \right\|,$$

$$R_5 = a \cdot (b - 1).$$

$$A_6 = \left\| \left(\delta - \frac{1}{a}\right)_i^j \left(\delta - \frac{1}{b}\right)_k^l \right\|,$$

$$R_6 = (a - 1)(b - 1).$$

$$A_7 = \left\| \left(\delta - \frac{1}{a}\right)_i^s \left(\delta - \frac{1}{a}\right)_t^j y_s y^t \right\| = \left\| \left[\left(\delta - \frac{1}{a}\right)^s y_s\right]_i \left[\left(\delta - \frac{1}{a}\right)_t y^t\right]^j \right\|,$$

$$R_7 = 1 \cdot 1 = 1.$$

The Matrix A_7 may be confusing at first sight. Note that each element, α_i^j, is a quadratic form in the y's. This form is of rank 1 and can be factored into two linear factors, one independent of j, the other independent of i.

To illustrate the application of these techniques to a fairly complicated problem, we shall construct and verify a design for the analysis of variance involving a regression line. It is known that sufficient conditions for such a design to be valid are:

1. The sum of the quadratic forms be equal to the sum of the sauares of the variables, and
2. The sum of the ranks of the forms be equal to the number of variables.

We shall use the first condition to set up our design. Thus,

$$
\begin{aligned}
\Sigma x_{ij}^2 &= [\delta\delta]_{ij}^{kl} x_{kl} x^{ij}, \\
&\equiv \left\{ \left[\delta\delta - \delta\frac{1}{b} - \frac{1}{a}\delta + \frac{1}{a}\frac{1}{b}\right]_{ij}^{kl} + \left[\frac{1}{a}\frac{1}{b}\right]_{ij}^{kl} \right. \\
&\quad + \left[\frac{1}{a}\delta - \frac{1}{a}\frac{1}{b}\right]_{ij}^{kl} + \left[\left(\delta - \frac{1}{a}\right)_s^k \left(\delta - \frac{1}{a}\right)_i^t y^s y_t \left(\frac{1}{a\sigma_y^2}\right)\left(\frac{1}{b}\right)_j^l\right] \\
&\quad \left. + \left[\left(\delta\frac{1}{b} - \frac{1}{a}\frac{1}{b}\right)_{ij}^{kl} - \left(\delta - \frac{1}{a}\right)_s^k \left(\delta - \frac{1}{a}\right)_i^t y^s y_t \left(\frac{1}{a\sigma_y^2}\right)\left(\frac{1}{b}\right)_j^l\right] \right\} x_{kl} x^{ij}.
\end{aligned} \tag{6}
$$

Rewriting this in the usual notation, we have for our tenative design,

$$
\begin{aligned}
\Sigma x_{ij}^2 = \Sigma[x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x}]^2 + \Sigma[\bar{x}]^2 + \Sigma[\bar{x}_{\cdot j} - \bar{x}]^2 \\
+ \Sigma[(r\sigma_x/\sigma_y)(y_i - \bar{y})]^2 + \Sigma[(\bar{x}_{i\cdot} - \bar{x}) - (r\sigma_x/\sigma_y)(y_i - \bar{y})]^2.
\end{aligned} \tag{7}
$$

In order to determine the corresponding equation for the ranks, we rewrite (6) in the form,

$$
\begin{aligned}
\Sigma x_{ij}^2 = \left\{ \left(\delta - \frac{1}{a}\right)_i^k \left(\delta - \frac{1}{b}\right)_j^l + \left(\frac{1}{a}\right)_i^k \left(\frac{1}{b}\right)_j^l + \left(\frac{1}{a}\right)_i^k \left(\delta - \frac{1}{b}\right)_j^l \right. \\
+ \left[\left(\delta - \frac{1}{a}\right)_s y^s\right]^k \left[\left(\delta - \frac{1}{a}\right)^t y_t\right]_i \left(\frac{1}{a\sigma_y^2}\right)\left(\frac{1}{b}\right)_j^l \\
\left. + \left[\left(\delta - \frac{1}{a}\right)_i^k - \left(\delta - \frac{1}{a}\right)_s^k \left(\delta - \frac{1}{a}\right)_i^t y^s y_t \left(\frac{1}{a\sigma_y^2}\right)\right]\left(\frac{1}{b}\right)_j^l \right\} x_{kl} x^{ij}.
\end{aligned} \tag{8}
$$

First we must determine the rank of the unfamiliar matrix,

$$
A_8 = \left\| \left(\delta - \frac{1}{a}\right)_i^j - \left(\delta - \frac{1}{a}\right)_s^j \left(\delta - \frac{1}{a}\right)_i^t y^s y_t / a\sigma_y^2 \right\|.
$$

We see that the rank of A_8 cannot be greater than $a - 2$ because two linear relations exist between the rows, namely,

$$
1^i \alpha_i^j = 0, \quad \text{since} \quad 1^i \left(\delta - \frac{1}{a}\right)_i^t = 0,
$$

$$
y^i \alpha_i^j = 0, \quad \text{since} \quad \left(\delta - \frac{1}{a}\right)_i^t y_t y^i = a\sigma_y^2.
$$

To show that the rank of A_8 cannot be less than $a - 2$, we subtract the elements of the first row from the corresponding elements of each of the last $a - 2$ rows, giving,

$$A_8 \sim \left\| \begin{array}{c|c|c} \alpha_i^1 & \alpha_i^2 & \alpha_i^j \\ \hline \alpha_i^1 - \alpha_1^1 & \dfrac{-\left(\delta - \dfrac{1}{a}\right)_s^2 y^s(\delta_i^t - \delta_1^t)y_t}{a\sigma_y^2} & \delta_i^j - \dfrac{\left(\delta - \dfrac{1}{a}\right)_s^j y^s(\delta_i^t - \delta_1^t)y_t}{a\sigma_y^2} \end{array} \right\| \begin{array}{l} i = 1, 2 \\ \\ i \neq 1, 2 \end{array}.$$

Multiplying each element of the second column by $-\left(\delta - \dfrac{1}{a}\right)_s^j y^s \Big/ \left(\delta - \dfrac{1}{a}\right)_s^2 y^s$ and adding the result to the corresponding element of the jth column for $j = 3, 4, \cdots a$, we see that the $(a - 2)$th order determinant in the lower right-hand corner becomes $| \delta_i^j |$ which is not equal to zero. Therefore the rank of A_8 must equal $a - 2$.

Referring to equation (8), we now write down the corresponding equation for ranks using the theorem on uncontracted products. Thus,

$$\begin{aligned} \Sigma \text{ Ranks} &= (a - 1)(b - 1) + (1)(1) + (1)(b - 1) + (1)(1)(1) + (a - 2)(1), \\ &= ab. \end{aligned}$$

Hence the quadratic forms in the right member of equation (7) are mutually independent and each, measured in units of the variance of the population, is distributed as is Chi-square with the appropriate number of degrees of freedom.