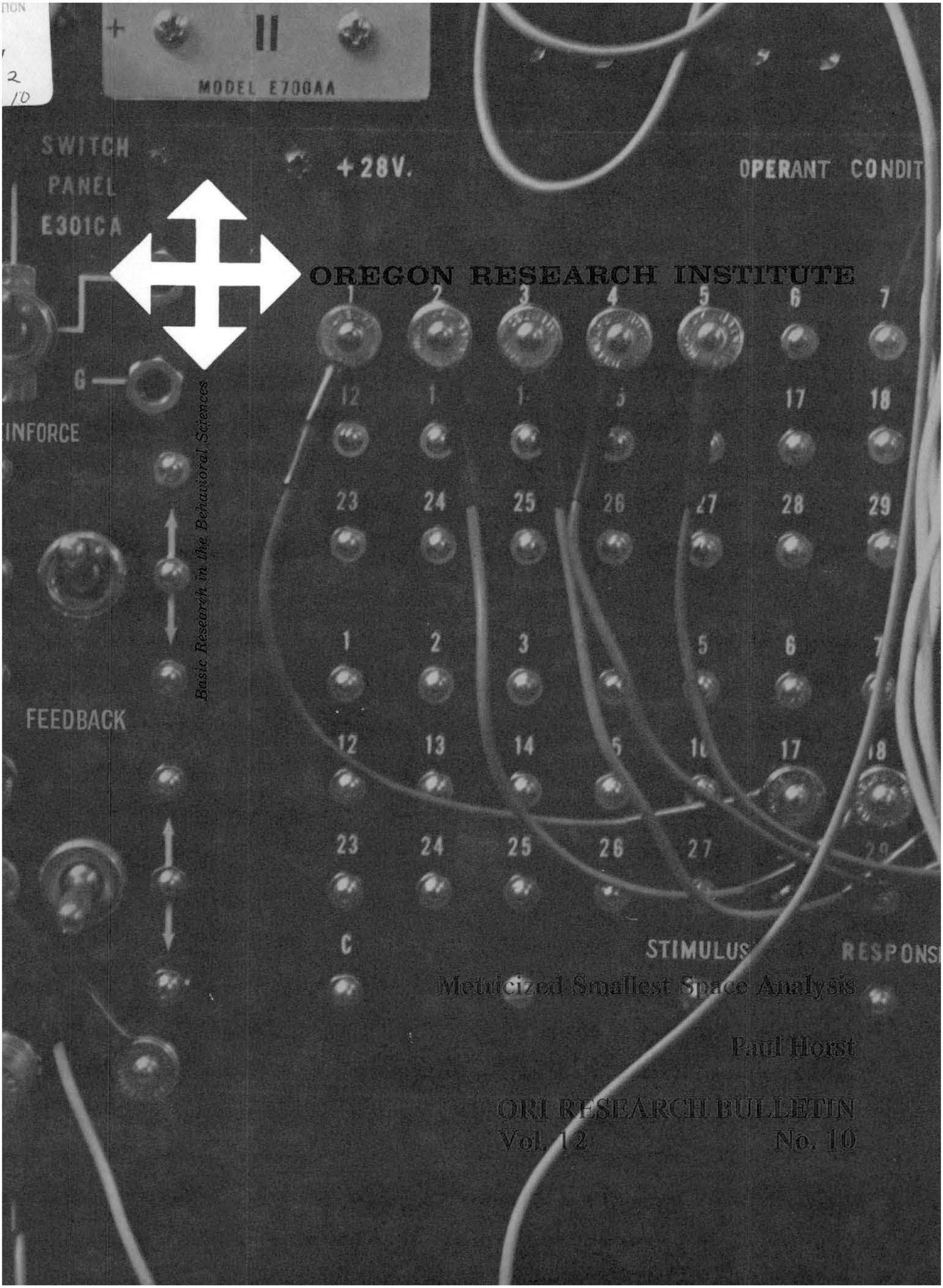

OREGON RESEARCH INSTITUTE

Basic Research in the Behavioral Sciences

Metricized Smallest Space Analysis

Paul Horst

ORI RESEARCH BULLETIN
Vol. 12 No. 10

OREGON RESEARCH INSTITUTE

METRICIZED SMALLEST SPACE ANALYSIS

Paul Horst

Oregon Research Institute

Research Bulletin
Vol. 12 No. 10
1972

ABSTRACT

The correlation-attenuating effects of arbitrary and fortuitous metrics in multivariate analysis procedures, together with previous attempts to circumvent these, are discussed. A new procedure utilizes side conditions imposed on reduced rank models and proposes the use of maximum covariance matrices in formulating these conditions. Approximations to the side conditions are developed and the procedures are applied to the binary simplex matrix. Expressions for maximum covariances as functions of relative frequency distributions are developed and special cases are considered.

METRICIZED SMALLEST SPACE ANALYSIS

I. The Fortuitous Metric Problem

A fundamental limitation of most, if not all, reduced rank approximations to data matrices is the fact that such reductions are dependent on variations in the moments of the distributions of the variables in a sample of entity vectors for a given set of attributes. Specifically, a correlation of unity can occur only if the frequency vectors are equal for both variables.

Much attention has been given to problems involving correlations among binary measures. These problems arise because a correlation of unity between two such variables is possible only if their means or P values are equal. If one wishes to construct a unidimensional test from a set of dichotomously-scored items which will yield more than dichotomous discrimination among entities, one must have a range of "difficulty" or P values for the items. But, if the set is to be considered as unidimensional, this means that the items must have maximum intercorrelations. Strictly speaking, if the set is unidimensional, the matrix of measures would by definition be of rank one. But, it is well known that a matrix of binary measures cannot exhibit other than a dichotomous dispersion of item difficulties and still be of unit rank. As a matter of fact, the rank of a binary data matrix can in no case be less than the number of distinct item difficulties in the set.

II. Attempted Solutions to the Problem

This dilemma has generated a number of discussions and proposals for the treatment of binary and other discrete measures. The work of Guttman (1955) on the simplex recognized the limitations of the traditional factor analytic and other multivariate analysis models when applied to binary measures. The later nonmetric approaches of Shepard (1962), Kruskal (1964), and Coombs (1964), and the smallest-space analysis procedures of Guttman (1968) and Lingoes (1968), have directly or indirectly attempted to circumvent the problems introduced by the attenuating or contaminating influence of disparate distributions among variables.

Carroll (1961) has proposed the application of tetrachoric correlations to a wide range of correlational data. In recent years, Horst (1964, 1965, 1968) has discussed the problem of homogeneity vs. dimensionality encountered with disparate distributions and proposed several procedures for handling binary measures. Bentler (1970) has pointed out the limitations of most of the proposed approaches and has suggested that some of these proposals (e.g., Horst, 1965) "pour out the baby with the bathwater." While the question of which is the baby and which the bathwater is clearly still under debate, Bentler has presented an insightful and penetrating critique of current approaches and their limitations. He suggested an approach, which he called "monotonicity analysis," in an effort to circumvent the limitations of previous methods. Unfortunately, his presentation does not follow a tightly structured algebraic development, and his procedure, like the tetrachoric approach, may well discard fragments of the baby along with the bathwater.

But whatever the shortcomings of methods proposed to date, such efforts are long overdue. Much of the work in psychological measurement suffers from a failure either to recognize the problem or to provide adequate solutions. It is the purpose of this report to suggest a general approach to reduced rank treatment of data contaminated by arbitrary or fortuitous metrics, including most data encountered in psychology and perhaps in other disciplines as well. The most immediate and urgent application of the method is in the measurement of personality by means of responses to statements. However, the general approach will be presented first and specific cases developed later.

III. Reduced Rank Approximation with Side Conditions

Suppose we let X be an $N \times n$ data matrix with N entities and n attributes. We shall consider first the problem of approximating a data matrix with one whose covariance matrix satisfies prespecified conditions.

Let the covariance matrix of X be given by

$$C_X = X' X \tag{3.1}$$

For the present, we shall not specify scales or origins of the variables in X . We let

$$U = Y A' \tag{3.2}$$

where

$$Y' Y = I \tag{3.3}$$

$$A A' = \sigma , \tag{3.4}$$

the widths of A and Y are less than that of X , U is the approximation to X, and σ is the specified covariance matrix of U . We write

$$X - U = \epsilon \tag{3.5}$$

We wish to minimize $\operatorname{tr} \epsilon \epsilon'$ with the constraint (3.3). Without loss of generality, we let

$$A = Q \Delta \tag{3.6}$$

where Δ is positive diagonal and

$$Q' Q = I \tag{3.7}$$

We let

$$\phi = \operatorname{tr} \epsilon \epsilon' \tag{3.8}$$

$$\psi = \phi + \operatorname{tr} Y' Y \lambda \tag{3.9}$$

where λ is a symmetric matrix of Lagrangian multipliers. From (3.1), (3.5), (3.8), and (3.9)

$$\psi = \operatorname{tr} (C - Y A' X - X A Y' + Y A'A Y' - Y \lambda Y') \tag{3.10}$$

We set

$$\frac{\partial \psi}{\partial Y'} = 0 \tag{3.11}$$

From (3.10) and (3.11)

$$X\ A = Y\ (A'\ A - \lambda) \tag{3.12}$$

Let

$$\gamma = (A'\ A - \lambda)^{-1} \tag{3.13}$$

From (3.12) and (3.13)

$$X\ A\ \gamma = Y \tag{3.14}$$

A necessary and sufficient condition that Y in (3.14) satisfies (3.3) is

$$\gamma = (A'\ C\ A)^{-\frac{1}{2}} \tag{3.15}$$

From (3.14) and (3.15)

$$Y = X\ A\ (A'\ C\ A)^{-\frac{1}{2}} \tag{3.16}$$

From (3.2) and (3.16)

$$U = X\ A\ (A'\ C\ A)^{-\frac{1}{2}}\ A' \tag{3.17}$$

More generally, it can be verified from (3.4), (3.6), and (3.17) that

$$U = X\ \sigma^{\frac{1}{2}}\ (\sigma^{\frac{1}{2}}\ C\ \sigma^{\frac{1}{2}})^{-\frac{1}{2}}\ \sigma^{\frac{1}{2}} \tag{3.18}$$

From (3.4) and (3.17), or from (3.18), it can be seen that

$$U'\ U = \sigma \tag{3.19}$$

IV. The Complete Decontaminated Covariance Matrix

We see then that the matrix U satisfies (3.19), the prespecified covariance matrix, and that with this constraint it is the least square approximation to the data matrix.

Next, we shall consider the matrix σ as it relates to the disparate distribution or fortuitous metric problem. Consider first the matrix C in equation 3.1. This is, of course, a covariance matrix. In particular, let us assume that X has been processed so that C is a correlation matrix. Let $f_1, - - , f_n$ be the frequency proportion vectors for the n columns of measures in X. It is known that the correlations in C cannot all be unity unless all f_i vectors are equal. But, as shown in Section 7, it is possible to find the maximum possible covariance between variables i and j for the vectors f_i and f_j. We can also, of course, find the maximum correlations, since the standard deviations corresponding to a maximum covariance are the same as for the sample.

Suppose now we let G be a matrix of these maximum correlations. We define

$$\sigma = C / G \tag{4.1}$$

where the division symbol on the right indicates elemental division. In general, it is clear from (4.1) that

$$|c_{ij}| \lesssim |\sigma_{ij}| \lesssim 1 \tag{4.2}$$

The inequalities in (4.2) hold for all cases in which f_i and f_j are distinct.

It will be recognized at once that, for dichotomous variables, σ in (4.1) is a matrix of the well known ratios of obtained phi coefficients to maximum phi coefficients. Such coefficients have come in for criticism by

Carroll (1961) and others for reasons that may not be completely convincing. Some may question why a ratio should be used rather than some other function which would give unity as an upper limit. Since the ratio is the most simple function of the two values which has an upper limit of unity, perhaps one might pardonably shift the burden of proof to those who suggest other functions.

V. The Lower Rank Approximation

In any case, it is well known that as the elements of a correlation matrix increase, so does the positive skewness of the distribution of its roots. It should also be intuitively obvious that, as this positive skewness increases, the lowest rank of a matrix which approximates the original with a specified degree of accuracy decreases.

Suppose therefore we let

$$\sigma = Q \Delta^2 Q' \tag{5.1}$$

$$\sigma_m = Q_m \Delta_m{}^2 Q_m' \tag{5.2}$$

where m indicates the m largest roots and corresponding vectors of σ. We may then write, instead of (3.18),

$$U_m = X \sigma_m{}^{\frac{1}{2}} (\sigma_m{}^{\frac{1}{2}} C \sigma_m{}^{\frac{1}{2}})^{-\frac{1}{2}} \sigma_m{}^{\frac{1}{2}} \tag{5.3}$$

Let us also indicate the basic structure of C by

$$C = q d^2 q' \tag{5.4}$$

and its rank m approximation by

$$C_m = q_m d_m{}^2 q_m' \tag{5.5}$$

Instead of (5.3), we then write

$$W_m = X\, q_m\, q_m' \tag{5.6}$$

We would then expect, in general, that

$$\mathrm{tr}\, U_m'\, U_m > \mathrm{tr}\, W_m'\, W_m \tag{5.7}$$

and that this inequality would increase as the disparities among the f_i increase.

VI. The Simplex Binary Matrix

In particular, the foregoing approach gives us a simple way out of the simplex dilemma for a binary simplex matrix encountered in the perfect Guttman scale. If X is a standardized perfect simplex matrix of binary scores, then the elements of the matrix C take their maximum value. Therefore

$$C = G \tag{6.1}$$

From (4.1) and (6.1)

$$\sigma = 1\, 1' \tag{6.2}$$

From (6.2)

$$\sigma^{\frac{1}{2}} = \frac{1\, 1'}{\sqrt{N}} \tag{6.3}$$

Substituting (6.3) in (3.18)

$$U = \frac{X\, 1\, 1'}{\sqrt{1' C\, 1}} \tag{6.4}$$

where, of course, U in (6.4) is a rank-one matrix.

VII. Maximum Covariance

Suppose we have given two measures for N entities on each of two attributes, X and Y, with proportional frequency vectors f_X and f_Y of length n_X and n_Y, respectively. We wish to determine the maximum covariance possible for the given vectors. Obviously, this maximum would occur if ordering entities by X would also order them by Y. Without loss of generality, ordering within an interval is arbitrary. We can therefore consider two N-high orthogonal binary matrices, L_X and L_Y, of width n_X and n_Y, respectively. The matrix L_X is an n_X-high diagonal supermatrix whose diagonal elements are unit vectors, each of length equal to N times the corresponding f_X element. L_Y is similarly defined for f_Y. Without loss of generality, we let the lowest scores for both X and Y be unity, and therefore the highest are n_X and n_Y, respectively.

We let

$$V_X' = (1 , - - , n_X) \tag{7.1}$$

$$V_Y' = (1 , - - , n_Y) \tag{7.2}$$

and X and Y be the N th order vectors of X and Y scores. From the definitions we have

$$X = L_X V_X \tag{7.3}$$

$$Y = L_Y V_Y \tag{7.4}$$

We let T_X and T_Y be two triangular matrices of order

n_X and n_Y , respectively, with all nonzero elements unity. Therefore, from (7.1) and (7.2), respectively, we have

$$T_X\, 1 = V_X \tag{7.5}$$

$$T_Y\, 1 = V_Y \tag{7.6}$$

Now the covariance between X and Y is given by

$$C_{XY} = \frac{X'\,Y}{N} - \frac{X'\,1}{N}\,\frac{1'\,Y}{N} \tag{7.7}$$

To find the maximum value of C_{XY} for given f_X and f_Y , we consider first

$$\frac{(X + Y)'\,(X + Y)}{N} = \frac{X'\,X + 2\,X'\,Y + Y'\,Y}{N} \tag{7.8}$$

From (7.8) we have

$$\frac{2\,X'\,Y}{N} = \frac{(X + Y)'\,(X + Y) - X'\,X - Y'\,Y}{N} \tag{7.9}$$

From (7.3) through (7.6) we have

$$X + Y = \begin{bmatrix} L_X \;, & L_Y \end{bmatrix} \begin{bmatrix} T_X & 0 \\ 0 & T_Y \end{bmatrix} \begin{bmatrix} 1_X \\ 1_Y \end{bmatrix} \tag{7.10}$$

Now let us indicate a permutation matrix by

$$\pi \equiv \begin{bmatrix} \pi_{XX} & \pi_{XY} \\ \pi_{YX} & \pi_{YY} \end{bmatrix} \tag{7.11}$$

which we shall define subsequently, where the subscripts indicate the orders of the matrix elements. Let

$$X + Y = S \tag{7.12}$$

From (7.10), (7.11), and (7.12)

$$S \equiv \begin{bmatrix} L_X , L_Y \end{bmatrix} \begin{bmatrix} T_X & 0 \\ 0 & T_Y \end{bmatrix} \begin{bmatrix} \pi_{XX} & \pi_{XY} \\ \pi_{YX} & \pi_{YY} \end{bmatrix} \begin{bmatrix} 1_X \\ 1_Y \end{bmatrix} \tag{7.13}$$

Let

$$L_X T_X = F_X \tag{7.14}$$

$$L_Y T_Y = F_Y \tag{7.15}$$

From the definitions, F_X and F_Y are both binary simplicial matrices with zeros in the upper right-hand corners. From (7.13), (7.14), and (7.15) we have

$$S \equiv \begin{bmatrix} F_X , F_Y \end{bmatrix} \begin{bmatrix} \pi_{XX} & \pi_{XY} \\ \pi_{YX} & \pi_{YY} \end{bmatrix} \begin{bmatrix} 1_X \\ 1_Y \end{bmatrix} \tag{7.16}$$

Consider now

$$\frac{1' (F_X , F_Y)}{N} = (P_X' , P_Y') \tag{7.17}$$

It can be seen that the right-side elements of (7.17) are cumulative frequency vectors of f_X and f_Y, respectively, with P_{X_1} and P_{Y_1} equal to unity.

Now let

$$P_S' \equiv \begin{bmatrix} P_X' , P_Y' \end{bmatrix} \begin{bmatrix} \pi_{XX} & \pi_{XY} \\ \pi_{YX} & \pi_{YY} \end{bmatrix} \tag{7.18}$$

where now π is defined such that

$$P_{S_i} \geqslant P_{S_{i+1}} \tag{7.19}$$

Let

$$F_S \equiv \begin{bmatrix} F_X , F_Y \end{bmatrix} \begin{bmatrix} \pi_{XX} & \pi_{XY} \\ \pi_{YX} & \pi_{YY} \end{bmatrix} \tag{7.20}$$

From the definitions, therefore, F_S is an $N \times (n_X + n_Y)$ binary simplicial matrix.

From (7.16) and (7.20) we have

$$\frac{S}{N} = \frac{F_S \, 1_S}{N} \tag{7.21}$$

where 1_S is of order $n_X + n_Y$. From (7.21)

$$\frac{S' \, S}{N} = \frac{1_S' \, F_S' \, F_S \, 1_S}{N} \tag{7.22}$$

But from (7.22) and (7.18)

$$\frac{S' \, S}{N} = \begin{bmatrix} P_1 + P_2 + P_3 + - - + P_{n_X+n_Y} \\ + P_2 + P_2 + P_3 + - - + P_{n_X+n_Y} \\ + P_3 + P_3 + P_3 + - - + P_{n_X+n_Y} \\ - \quad - \quad - \quad - \quad - \\ + P_{n_X+n_Y} \quad - \quad - \quad P_{n_X+n_Y} \end{bmatrix} \tag{7.23}$$

Let

$$V_S' = (V_X' \, , \, n_X \, 1 + V_Y') \tag{7.24}$$

From (7.23) and (7.24)

$$\frac{S' \, S}{N} = 2 \, V_S' \, P_S - 1' \, P_S \tag{7.25}$$

It can also be shown that

$$\frac{X' \, X}{N} = 2 \, V_X' \, P_X - 1' \, P_X \tag{7.26}$$

$$\frac{Y' \, Y}{N} = 2 \, V_Y' \, P_Y - 1' \, P_Y \tag{7.27}$$

From (7.9), (7.12), (7.25), (7.26), and (7.27)

$$\frac{2\ X'Y}{N} = 2\ V_S'P_S - 1'P_S - 2\ V_X'P_X + 1'P_X - 2\ V_Y'P_Y + 1'P_Y \tag{7.28}$$

But

$$1'\ P_S = 1'\ P_X + 1'\ P_Y \tag{7.29}$$

From (7.28) and (7.29)

$$\frac{X'\ Y}{N} = V_S'\ P_S - V_X'\ P_X - V_Y'\ P_Y \tag{7.30}$$

From (7.7) and (7.30)

$$C_{XY} = V_S'\ P_S - V_X'\ P_X - V_Y'\ P_Y - P_X'\ 1\ 1'\ P_Y \tag{7.31}$$

VIII. Change of Origin from "1" to "0"

Let

$$v_S = V_S - 2\cdot 1 \tag{8.1}$$

$$v_X = V_X - 1 \tag{8.2}$$

$$v_Y = V_Y - 1 \tag{8.3}$$

By definition

$$P_{S_1} = P_{S_2} = P_{X_1} = P_{Y_2} = 1 \tag{8.4}$$

Let

ρ_S be obtained by deleting the first two elements of P_S , and ρ_X and ρ_Y by deleting the first element of P_X and P_Y , respectively. From (7.31) through (8.3)

$$C_{XY} = v_S'P_S + 2\ 1'P_S - v_X'P_X - 1'P_X - v_Y'P_Y + 1'P_Y - P_X'1\ 1P_Y' \tag{8.5}$$

But from (7.29) and (8.4)

$$C_{XY} = v_S' P_S + 1' P_S - v_X' P_X - v_Y' P_Y - P_X' 1 1' P_Y \tag{8.6}$$

From the definitions of the ρ's, and from (8.4) and (8.6)

$$v_S' P_S = v_S' \rho_S - 1 \tag{8.7}$$

$$1' P_S = 1' \rho_S + 2 \tag{8.8}$$

$$v_X' P_X = v_X' \rho_X \tag{8.9}$$

$$v_Y' P_Y = v_Y' \rho_Y \tag{8.10}$$

$$P_X' 1 = \rho_X' 1 + 1 \tag{8.11}$$

$$1' P_Y = 1' \rho_Y + 1 \tag{8.12}$$

Redefine

$$v_S' = (1 \ , \ - -, (n_X + n_Y - 2)) \tag{8.13}$$

$$v_X' = (1 \ , \ - -, (n_X - 1)) \tag{8.14}$$

$$v_Y' = (1 \ , \ - -, (n_Y - 1)) \tag{8.15}$$

From (7.29) and (8.6) through (8.15)

$$C_{XY} = v_S' \rho_S - v_X' \rho_X - v_Y' \rho_Y - \rho_X' 1 1' \rho_Y \tag{8.16}$$

IX. Computational Procedure

To calculate maximum C_{XY} in (8.16), we proceed as follows:

1. Calculate f_X and f_Y .
2. Calculate both ρ_X and ρ_Y by

$$\rho_1 = 1 - f_1 \tag{9.1}$$

$$\rho_i = \rho_{i-1} - f_i \qquad (9.2)$$

$$\rho_{n-1} = f_n \qquad (9.3)$$

3. Calculate $\rho_X' 1$ and $\rho_Y' 1$.

4. Combine the elements of ρ_X and ρ_Y to form ρ_S so that

$$\rho_{S_i} \geqslant \rho_{S_{i+1}} \qquad (9.4)$$

5. Calculate $v_S' \rho_S$, $v_X' \rho_X$, and $v_Y' \rho_Y$.

6. Calculate C_{XY} from the terms calculated in instructions 4 and 6 by means of equation 8.16.

Essentially, this method for calculating the maximum value of a covariance for given proportional frequency vectors was given without proof by Carroll (1961).

X. Special Cases

We shall next consider three important special cases of equation 8.16.

First, we consider the case when $X = Y$. From (8.13) it can be shown that

$$v_S' \rho_S = 4\ v_X' \rho_X - 1' \rho_X \qquad (10.1)$$

From (8.16) and (10.1)

$$C_{XX} = 2\ v_X' \rho_X - 1' \rho_X\ (1 + 1' \rho_X) \qquad (10.2)$$

Obviously, equation 10.2 gives the variance of X as a function of the cumulative frequency proportions.

Next, we consider the case where $n_X = 2$ and therefore

$$v_X = 1 \tag{10.3}$$

This is, of course, the case of a point biserial covariance. Suppose that

$$\rho_{Y_{k-1}} \geqslant \rho_X \geqslant \rho_{Y_k} \tag{10.4}$$

and let

$$_k\rho_Y' = (\rho_{Y_{k+1}} , - - \rho_{Y_{n_Y - 1}}) \tag{10.5}$$

It can readily be proved from (10.4) and (10.5) that $v_S' \rho_S$ in (8.16) is given by

$$v_S' \rho_S = v_Y' \rho_Y + 1' \, _k\rho_Y + {}_k\rho_X \tag{10.6}$$

Then from (8.16), (10.3), and (10.6) we have

$$C_{XY} = 1' \, _k\rho_Y + \rho_X (k - 1 - 1' \rho_Y) \tag{10.7}$$

To my knowledge, formula 10.7 for the maximum point biserial, given f_X and f_Y , has not been published.

Finally, we shall assume

$$v_X = 1 \tag{10.8}$$

$$v_Y = 1 \tag{10.9}$$

This is the case of two dichotomous variables. Because of (10.8) and (10.9), ρ_X and ρ_Y are scalars. Without loss of generality, we shall assume

$$\rho_X \geqslant \rho_Y \tag{10.10}$$

It can readily be shown then that

$$v_S' \rho_S = \rho_X + 2 \rho_Y \tag{10.11}$$

$$v_X' \rho_X = \rho_X \qquad (10.12)$$

$$v_Y' \rho_Y = \rho_Y \qquad (10.13)$$

$$\rho_X' 1 = \rho_X \qquad (10.14)$$

$$\rho_Y' 1 = \rho_Y \qquad (10.15)$$

Substituting equations 10.11 through 10.15 in equation 8.16, we get

$$c_{XY} = \rho_Y - \rho_X \rho_Y \qquad (10.16)$$

Equation 10.16 is the well known expression for the maximum possible covariance between two dichotomous variables when equation 10.10 is satisfied (Horst, 1966). It is this formula that should be used to calculate the G and σ matrices of equation 4.1 when doing a factor analysis on dichotomously scored personality items.

XI. The Case of Negative Correlations

In the foregoing analysis, there is the implicit assumption that all of the intercorrelations are positive. Let us assume that variables have been reflected so that the sum of each row of the correlation matrix is greater than unity. This can always be done (see Horst, 1965, p. 131 ff). Then for each remaining negative element in the C matrix of equation 4.1, we should determine whether the largest possible value is equal to or greater than the absolute value of the largest possible negative value. Whichever is greater in absolute value should be used as a positive divisor in getting the corresponding element in σ of equation 4.1.

Carroll (1961) has stated that, to get the maximum negative value, one simply reverses the orientation of one of the distributions. Although he presents no algebraic proof, the procedure appears intuitively obvious so that we shall not attempt one here.

However, it may be useful to consider the special case of maximum negative covariances for the case of binary variables X and Y. Assume that

$$P_X \geqslant P_Y \tag{11.1}$$

Then the maximum covariance is well known to be

$$_{max}C_{XY} = (1 - P_X)\, P_Y \tag{11.2}$$

It can readily be shown that if

$$P_X + P_Y > 1 \tag{11.3}$$

then the absolute minimum value G_{XY} can take is

$$\left| _{min}G_{XY} \right| = (1 - P_X)\, (1 - P_Y) \tag{11.4}$$

If

$$P_X + P_Y \leqslant 1 \tag{11.5}$$

then we have

$$\left| _{min}G_{XY} \right| = P_X\, P_Y \tag{11.6}$$

Therefore, for each negative value in C remaining after optimal reflection, we proceed as follows:

1. If (11.3) holds, calculate (11.4).

2. If (11.5) holds, calculate (11.6).

3. Find the larger of (11.2) and either instruction 1 or 2 above and call this G_{XY} .

4. These G_{XY} values are the ones to be used in the G of (4.1) for negative values of C_{XY} .

XII. Number of Factors and Specificities

In Section V we have considered the lower m-rank approximation to the σ matrix without specifying m . Here, of course, we encounter the everpresent problem of how many factors to solve for. But before suggesting criteria for this purpose, it is important to note that the σ matrix as defined in Section IV may not even be Grammian. We propose therefore that a dual criterion be adopted for determining m in (5.2):

1. In any case all the elements in $\Delta_m{}^2$ shall be greater than unity.

2. In no case shall $\operatorname{tr}(\sigma - \sigma_m)$ be less than zero.

The first of these criteria is the familiar Kaiser criterion of eigenvalues greater than unity. The second means that the sum of the roots in the residual matrix $(\sigma - \sigma_m)$ is not less than zero.

It will be noted that throughout this presentation the communality-specificity issue has not been raised. The model implied by the foregoing development does not assume specific non-error variance. Indeed, for the special case of the error-free binary simplex matrix, the conventional specific factor models such as "maximum likelihood," "minres," and "alpha" cannot yield rank-one matrices as in Section VI.

It should also be noted that the rank-m solutions for the Y and A matrices of (3.2) can be followed with any of the available simple structure transformation procedures.

XIII. Use of Tetrachoric Correlations

It should be pointed out that the alternative proposal of Carroll (1961) to use tetrachoric correlations in instead of the ratios indicated in (4.1), which he does not favor, could be equally well utilized in Section III, since no restrictions are implied except that σ is real and symmetric. Such a matrix of tetrachoric coefficients is not necessarily Grammian, just as the matrix of ratios need not be Grammian. One might argue that the arbitrariness of the ratio functions is no more objectionable than the assumption of underlying multi-normality in the tetrachoric functions.

The former has two possible advantages:

1. Computationally, it appears to be simpler and more straightforward.

2. It is unique, whereas the tetrachoric functions depend on the selection of cutting points for dichotomization in the case of more than two-valued variables.

One can, of course, as Carroll (1961) suggests, select the points as close to a 50-50 split as possible.

In any case, further theoretical and experimental work is required to determine which of the two functions is better, and indeed whether our proposals are adequate for most data.

REFERENCES

Bentler, P. M. A comparison of monotonicity analysis with factor analysis, Educational and Psychological Measurement, 1970, 30, 241-250.

Carroll, J. B. The nature of data, or how to choose a correlation coefficient, Psychometrika, 1961, 62, 347-372.

Coombs, C. H. A theory of data. New York: Wiley, 1964.

Guttman, L. An additive metric from all the principal components of a perfect scale, British Journal of Statistical Psychology, 1955, 8, 17-24.

Guttman, L. A general nonmetric technique for finding the smallest coordinate space for a configuration of points, Psychometrika, 1968, 33, 469-506.

Horst, P. Matrix factoring and test theory. In N. Frederiksen and H. Gulliksen (Eds.), Contributions to mathematical psychology. New York: Holt, 1964, 129-140.

Horst, P. Matrix algebra for social scientists. New York: Holt, 1965.

Horst, P. Factor analysis of data matrices. New York: Holt, 1966.

Horst, P. Personality: measurement of dimensions. San Francisco: Jossey-Bass, 1968.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika, 1964, 29, 115-129.

Lingoes, J. C. The multivariate analysis of qualitative data, Multivariate Behavioral Research, 1968, 3, 61-94.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function, I, II, *Psychometrika*, 1962, 27, 125-140, 219-246.