FACTOR ANALYSIS WITH EXTERNAL CRITERIA: APPLICATION OF ANALYSIS OF VARIANCE AND REGRESSION ANALYSIS TECHNIQUE TO FACTOR ANALYSIS

HARUO YANAI¹

University of Tokyo

We propose a new technique of factor analysis with external criteria together with its computational procedure.

First of all, we shall discuss some relationships between the principle of analysis of variance and that of regression analysis by introducing the mathematical concept of *projection*. Secondly, we shall show that two kinds of factors could be extracted, and that one of them would be closely related to the informations concerning the predetermined external criteria and another one would be independent of them. Thirdly, we shall show that our proposal might give a satisfactory answer to psychological interpretation of factors without assuming rotation of factors and the estimation of communalities. Finally, two numerical examples will be presented.

INTRODUCTION AND PROBLEMS

By the rapid development of electronic computer, it has been quite easy to have factor analytic studies based on a great number of variables. And, as for factor rotation, many plausible machine methods have been devised and they seem to give good interpretation. In spite of these proposals, however, it seems not always to assure invariance of interpretation of factors. The present author thinks that the trouble comes from the fact that the obtained results may be contaminated in actual data, and thinks that it comes from the fact that factor analyst do not make difference between the informations which

may be gotten before-hand and the unknown ones. And, they do not always make use of the already-known informations concerning subjects even when they are available at hand.

Gulliksen (1950), in his utopian theory of mental test scores, assumed that the test scores should be decomposed into two components which consist of true and error scores. By criticizing the theory, however, Cronbach (1951), and Cronbach and Cleser (1965) proposed the idea that the scores may have the possibility to be divided into more components than Gulliksen assumed, and that useful informations may still contain in the error score. And, independently of Cronbach, Tukey (1962) suggested the possibility of extending a similar kind of idea to multidimensional analysis and made the following suggestions.

When we perform factor analysis in usual sense, we are not always concerned with ruling out the informations which have been gotten before analysis starts. For example, even when we have before-hand the informations such as sex dif-

¹ The author thanks Mr. S. Kashiwagi of Japan National Railroads and Dr. H. Ikeda of Rikkyo University for their critical readings.

Dr. L. J. Cronbach of Stanford University gave us lecture concerning test theory in the University of Tokyo 1967 to 1968. The author was greatly impressed by his concept of "generalizability". And, I would like to extend my sincere gratitude to Dr. Cronbach who made valuable comments on this work.

ference and so on it is uncustomary to perform factorization based on the residual correlational matrix from which the already-known informations are taken off. However, Tukey suggested the importance of factorization of residual data and emphasized the urgent need to develop a new type of factor analysis following the principle as mentioned above. And, at the same time, he pointed out that the proposal would be a pioneering work in that it might seek out novelty in data analysis. It would be here noted, however, he did not suggest actual procedure for computation. Following the suggestion, we propose a new technique of factor analysis with external criteria, the principles of which may be said to be an application of analysis of variance and regression analysis to factor analysis.

Although a great number of researches based on either analysis of variance or factor analysis have been reported, researches based on the multidimensional analytic technique combining both seem to be very few except for the works by Burt (1966), Jones (1966) and Gollob (1969). Our proposal may differ from these kinds of studies in the sense that they did not combine simultaneously both analysis of variance and regression analysis with conventional factor analysis.

First of all we shall herein try to discuss some relationships between the principle of analysis of variance and that of regression analysis by introducing the mathematical concept of *projection* and then try to propose a new technique of factor analysis together with computational procedure which consists of four parts. Secondly, we shall show that two kinds of factors could be extracted in our own method, and that one of them would be closely related to the informations of the predetermined external criteria and another one would be independent of the informations related to the external criteria. And, we shall show the possibility to have the factors free from the already-known

informations, and shall show that the psychological meanings of the obtained factors may differ from the ones in conventional factor analysis. Finally, we shall show that our proposals might give a satisfactory answer to psychological interpretations of factors without assuming rotation of factors and the estimation of communalities, and that assertion above will be valid in the numerical examples applying the principle of our factor analysis to the construction of some vocational aptitude scales.

APPLICATION OF REGRESSION ANALYSIS TO FACTOR ANALYSIS

Let $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_m)$ be two sets of N dimensional vectors which we will designate as the elements of matrices X and Y . We will assume that the elements of vectors are represented in the deviation from such that the sum of the elements should be zero. We call Y predictors and X criteria in this paper.

Following the technique of regression analysis based on the principle of least squares, the vector (we omit the index i here) is decomposed into two components y_x and y_{x^-} . Then we have an equation

$$y = y_x + y_{x^-} = X(X'X)^{-1}X'y + (I_N - X(X'X)^{-1}X')y \quad (1)$$

where the elements of y_x express the components predictable perfectly by the elements of X and the elements of y_{x^-} express the components independent of X . I_N denotes N dimensional identity matrix.

Let

$$E_X = X(X'X)^{-1}X' \quad (2)$$

then the square matrix E_X satisfies the following two conditions

$$E_X' = E_X \quad \text{and} \quad E_X^2 = E_X \quad (3)$$

which is called the projection in mathematics, i.e., symmetric and idempotent.

We give one more definition such that

$$E_X^- = I_N - E_X \quad (4)$$

From (3) and (4), we get

$$E_X E_X^- = E_X (I_N - E_X) = E_X - E_X^2 = 0 \quad (5)$$

From this relation, y_X and y_X^- are easily proved to be orthogonal to each other. Taking the inner product of y as shown in (1), we get

$$\begin{aligned} (y, y) &= (y_x, y_x) + (y_x, y_x^-) \\ &\quad + (y_x^-, y_x) + (y_x^-, y_x^-) \\ &= (y_x, y_x) + (y_x^-, y_x^-) \end{aligned} \quad (6)$$

Division of both sides of (6) by N yields

$$\sigma y^2 = \sigma y_x^2 + \sigma y_x^{-2}. \quad (7)$$

This shows that the variance of y is decomposed into two components σy_x^2 and σy_x^{-2} , and the former term is predictable from X while the latter is not predictable.

By considering the case where all the predictor variables $Y = (y_1, y_2, \dots, y_m)$ exist, (1) is replaced by the following formula.

$$\begin{aligned} Y &= Y_X + Y_X^- = X(X'X)^{-1}X'Y \\ &\quad + (I_N - X(X'X)^{-1}X')Y \end{aligned} \quad (8)$$

where Y_X and Y_X^- show the matrices containing vectors

$$y_{1X}, y_{2X}, \dots, y_{mX}, y_{1X}^-, y_{2X}^-, \dots, y_{mX}^-$$

respectively.

When we premultiply (2) by Y' , we have the following decomposition of the covariance matrix C_{YY} .

$$\begin{aligned} C_{YY} &= \frac{1}{N} Y'Y = \frac{1}{N} Y'X(X'X)^{-1}X'Y \\ &\quad + \frac{1}{N} (Y'Y - Y'X(X'X)^{-1}X'Y) \\ &= C_{YX}C_{XX}^{-1}C_{XY} \\ &\quad + (C_{YY} - C_{YX}C_{XX}^{-1}C_{XY}) \end{aligned} \quad (8')$$

where C_{YX} and C_{XX} denote the covariance matrices of X and Y .

Let's define the first term and second term in (8) to be

$$C_{YY.X} = C_{YX}C_{XX}^{-1}C_{XY},$$

$$\text{and } C_{YY.X}^- = C_{YY} - C_{YX}C_{XX}^{-1}C_{XY} \quad (9)$$

Then we can perform factorization of C_{YY} by analyzing the matrices $C_{YX.X}$ and $C_{YY.X}^-$ separately rather than analyzing C_{YY} directly.

The factors extracted from $C_{YY.X}^-$ may be said to be free from the criterion variables, while the factors extracted from $C_{YX.X}$ by means of this technique are closely related to the criterion variables. Here it should be noted that, by factorization or factor analysis, we mean matrix decomposition which involves obtaining eigenvalues and eigenvectors. Communalities estimation might of course be considered in our method. However, we do not treat this problem here.

RELATIONSHIP BETWEEN ANALYSIS OF VARIANCE AND REGRESSION ANALYSIS

Let's consider the problem by making use of the projection E_X which we defined in (2). We here define a dummy matrix G_A as shown in Table 1. Each row element of this matrix consists of only one "1" and "0" otherwise.

By substituting the dummy matrix G_A into X in (4), the projection E_X can be represented as follows.

$$E_A = G_A(G_A'G_A)^{-1}G_A' \quad (10)$$

Expanding the middle factor of the right side, we get

$$(G_A'G_A)^{-1} = \begin{bmatrix} \frac{1}{N_1} & & & 0 \\ & \frac{1}{N_2} & & \\ & & \frac{1}{N_3} & \\ 0 & & & \frac{1}{N_m} \end{bmatrix} \quad (11)$$

where N_i denotes the number of the subjects belonging to the class i .

Let's define

TABLE 1
Dummy matrix representing class of subjects

	1	2	3	...	m
$G_A =$	1	0	0	...	0
	1	0	0	...	0
	0	1	0	...	0
	0	1	0	...	0
	0	0	1	...	0
					0
					
	0	0	0	...	1
	0	0	0	...	1
					

$$y = \begin{bmatrix} y_{11} - \bar{y} \\ y_{12} - \bar{y} \\ \dots \\ y_{iN} - \bar{y} \\ y_{21} - \bar{y} \\ y_{22} - \bar{y} \\ \dots \\ y_{m1} - \bar{y} \\ \dots \\ y_{mN} - \bar{y} \end{bmatrix} \quad (12)$$

By premultiplying y by E_A and expanding, we get

$$E_A y = G_A (G_A' G_A)^{-1} G_A' y = \begin{bmatrix} \bar{y}_1 - \bar{y} \\ \bar{y}_1 - \bar{y} \\ \dots \\ \bar{y}_2 - \bar{y} \\ \bar{y}_2 - \bar{y} \\ \dots \\ \bar{y}_m - \bar{y} \\ \dots \\ \bar{y}_m - \bar{y} \end{bmatrix} \quad (13)$$

where y_j denotes the mean score of y in the class j . Each element shows that the deviation score of the class mean from the

total mean. Then we have

$$E_A^- y = (I_N - E_A) y = y - E_A y$$

$$= \begin{bmatrix} y_{11} - \bar{y}_1 \\ y_{21} - \bar{y}_1 \\ \dots \\ \dots \\ y_{m1} - \bar{y}_m \\ y_{m2} - \bar{y}_m \\ \dots \\ y_{mN} - \bar{y}_m \end{bmatrix} \quad (14)$$

This shows the residual error term specific to individual difference.

Let y_A and y_{Ae} be defined to be $E_A y$ and $E_A^- y$, then y can be decomposed as follows.

$$y = E_A y + E_A^- y = y_A + y_{Ae} \quad (15)$$

As $E_A E_A^- = 0$, y_A and y_{Ae} become to be orthogonal to each other. This is shown as

$$(y_A, y_{Ae}) = y_A' y_{Ae} = 0. \quad (16)$$

By taking the inner product of y , we get the decomposition of the variance of y with respect to the factor level of A , which is equivalent to the one way analysis of variance technique. Then we have

$$\sigma y^2 = \sigma y_A^2 + \sigma y_{Ae}^2 \quad (17)$$

where σy_A^2 denotes the between-class variance and σy_{Ae}^2 denotes the residual error variance. As is the case with regression analysis, let's consider the case where multivariate predictor variables exist. In this case (15) is replaced by

$$Y = E_A Y + E_A^- Y \quad (18)$$

Then, by premultiplying Y by Y' , we get

$$Y' Y = Y' E_A Y + Y' E_A^- Y \quad (19)$$

Devision of (19) by N yields

$$C_{YY} = C_A + C_{Ae} \quad (20)$$

where C_A and C_{Ae} denote the covariance matrices corresponding to the between-class variance and the residual variance respectively.

Let D_C^2 be the diagonal matrix with the elements of the total variance $\sigma y_i^2 (1 \leq i \leq m)$. Pre- and postmultiplication of both sides of (20) yield the decomposition of the correlation matrix R_{YY} . This shows

$$R_{YY} = R_A + R_{Ae} \quad (21)$$

where

$$R_{YY} = D_C^{-1} C_{YY} D_C^{-1}, \quad R_A = D_C^{-1} C_A D_C^{-1}$$

and $R_{Ae} = D_C^{-1} C_{Ae} D_C^{-1}.$

In case of two way analysis of variance, the residual term y_{Ae} can be decomposed by the additional information such that the subjects classified with respect to the factor A can be classified with respect to the another factor B .

Without further details of derivation, we show that vector y can be decomposed as

$$y = y_A + y_B + y_{AB} + y_{ABe} \quad (22)$$

where y_B denotes the component specific to B and y_{AB} denotes the interaction effect between A and B , and where y_{ABe} denotes the residual term. In multivariate form, (22) is replaced by

$$Y = Y_A + Y_B + Y_{AB} + Y_{ABe}. \quad (23)$$

By the orthogonality condition² of the four terms shown in (23), the following decomposition of the covariance matrices can be derived.

$$C = C_A + C_B + C_{AB} + C_{ABe} \quad (24)$$

In correlational form, (24) is replaced by

$$R = R_A + R_B + R_{AB} + R_{ABe}. \quad (25)$$

In case of three way analysis of variance,

² Orthogonality condition of the four terms holds if only $N_{j..} N_{.li} = N N_{jli}$ where $N_{j.}$, $N_{.i}$ and N_{ji} denote the numbers of the subjects belonging to the class A_j , B_i and $A_j B_i$.

(24) and (25) are replaced by

$$C = C_A + C_B + C_C + C_{AB} + C_{BC} + C_{AC} + C_{ABC} + C_{ABCe} \quad (26)$$

and

$$R = R_A + R_B + R_C + R_{AB} + R_{BC} + R_{AC} + R_{ABC} + R_{ABCe} \quad (27)$$

where A , B and C denote the classification factors such as sex and vocation and so on. And, AB , BC and AC denote the interaction effects. ABe and $ABCe$ denote the residual term closely related to the individual difference free from these predetermined factors.

In conventional factor analysis, the correlation matrix R is analyzed directly. However, in our analysis R_A , R_B , ..., and R_{ABCe} are analyzed separately by the principal component method. As is the case with regression analysis, factors extracted from R_A , R_B , ..., and R_{ABC} are closely related to the predetermined level of factors while the factors extracted from R_{Ae} , R_{ABe} and R_{ABCe} are free from the predetermined ones.

In case three sets of variables are available

Let Y , X_1 and X_2 be three sets of variables, each of which is expressed in matrix notation. Projections constructed by the matrices X_1 and X_2 are defined as

$$\left. \begin{aligned} E_{X_1} &= X_1(X_1'X_1)^{-1}X_1' \\ E_{X_2} &= X_2(X_2'X_2)^{-1}X_2' \end{aligned} \right\} \quad (28)$$

and

By substituting X_1 into X in (2), we get

$$Y = Y_{X_1} + Y_{X_1}^- = E_{X_1}Y + E_{X_1}^-Y. \quad (29)$$

As the next step, we make $E_{X_1}Y$ and $E_{X_1}^-Y$ to be two sets of predictor variables, and X_2 to be criterion variables. Following the same line of derivations to get (2), we can decompose $E_{X_1}Y$ and $E_{X_1}^-Y$ as follows.

$$E_{X_1}Y = E_{X_2}(E_{X_1}Y) + E_{X_2}^-(E_{X_1}Y) \quad (30)$$

Fig. 1. Vector Representation of (31) and (32).

$$E_{X_1}^- Y = E_{X_2} (E_{X_1}^- Y) + E_{X_2}^- (E_{X_1}^- Y) \quad (31)$$

Vector representation of (30) and (31) is shown in Fig. 1.

As $E_{X_2} E_{X_2}^- = 0$, multiplication of (30) and (31) by the transpose of $E_{X_1} Y$ and $E_{X_1}^- Y$ respectively yields

$$C_{Y Y \cdot X_1} = \frac{1}{N} Y' E_{X_1} E_{X_2} E_{X_1} Y + \frac{1}{N} Y' E_{X_1} E_{X_2}^- E_{X_1} Y \quad (32)$$

$$C_{Y Y \cdot X_1}^- = \frac{1}{N} Y' E_{X_1}^- E_{X_2} E_{X_1}^- Y + \frac{1}{N} Y' E_{X_1}^- E_{X_2}^- E_{X_1}^- Y \quad (33)$$

In order to expand the both terms of (32) and (33), we set

$$C_{Y Y \cdot X_1 X_2} = \frac{1}{N} Y' E_{X_1} E_{X_2} E_{X_1} Y = C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2} C_{X_2 X_2}^{-1} \times C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 Y} \quad (34)$$

$$C_{Y Y \cdot X_1 X_2}^- = \frac{1}{N} Y' E_{X_1}^- E_{X_2}^- E_{X_1}^- Y = \frac{1}{N} Y' E_{X_1}^- Y - \frac{1}{N} Y' E_{X_1}^- E_{X_2} E_{X_1}^- Y = C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 Y} - C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2} C_{X_2 X_2}^{-1} \times C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 Y} \quad (35)$$

$$C_{Y Y \cdot X_1}^- X_2 = \frac{1}{N} Y' E_{X_1}^- E_{X_2}^- E_{X_1}^- Y = \frac{1}{N} Y' (E_{X_2}^- - E_{X_1}^- E_{X_2} - E_{X_2} E_{X_1})$$

$$+ E_{X_1} E_{X_2} E_{X_1}) Y = C_{Y X_2} C_{X_2 X_2}^{-1} C_{X_2 Y} - C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2} C_{X_2 X_2}^{-1} C_{X_2 Y} - C_{Y X_2} C_{X_2 X_2}^{-1} C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 Y} + C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2} C_{X_2 X_2}^{-1} \times C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 Y} \quad (36)$$

and

$$C_{Y Y \cdot X_1}^- X_2^- = \frac{1}{N} Y' E_{X_1}^- E_{X_2}^- E_{X_1}^- Y = \frac{1}{N} Y' E_{X_1}^- Y - \frac{1}{N} Y' E_{X_1}^- E_{X_2} E_{X_1}^- Y = C_{Y Y} - C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 Y} - C_{Y Y \cdot X_1}^- X_2 \quad (37)$$

From (32), (33), (34), (35), (36) and (37), the covariance matrix $C_{Y Y}$ can be decomposed into four component matrices. This shows

$$C_{Y Y} = C_{Y Y \cdot X_1} + C_{Y Y \cdot X_1}^- = C_{Y Y \cdot X_1 X_2} + C_{Y Y \cdot X_1 X_2}^- + C_{Y Y \cdot X_1}^- X_2 + C_{Y Y \cdot X_1}^- X_2^- \quad (38)$$

Application of the principal component method to each of the component matrices produces the factors closely related to four combinations of variables sets $X_1 X_2$, $X_1 X_2^-$, $X_1^- X_2$ and $X_1^- X_2^-$.

Another interpretation of (38) lies in that the total variances of the variables $y_i (1 \leq i \leq m)$ can be partitioned into four components, since the trace of the component matrix is equal to the sum of the variance of y_i which is explained by the criterion variables sets X_1 and X_2 . The factors extracted from $C_{Y Y \cdot X_1 X_2}$ and $C_{Y Y \cdot X_1 X_2}^-$ are mutually independent and the factors from $C_{Y Y \cdot X_1}^- X_2$ and $C_{Y Y \cdot X_1}^- X_2^-$ are also independent. However, it should be noted that the formers are not independent of the latters, since the orthogonality condition of E_{X_1} and E_{X_2} does not hold in general.

So we try to find the variables set orthogonal to the second variables set

X_1 . For this purpose, we try to eliminate the predictable part of X_2 on X_1 by premultiplying X_2 by $E_{X_1}^-$. Then we get

$$Z = E_{X_1}^- X_2 = X_2 - E_{X_1} X_2.$$

This variables set Z is easily proved to be independent of X_1 since

$$\begin{aligned} X_1' Z &= X_1' X_2 - X_1' E_{X_1} X_2 \\ &= X_1' X_2 - X_1' X_1 (X_1' X_1)^{-1} X_1' X_2 = 0. \end{aligned}$$

From this relation, we get

$$\begin{aligned} E_{X_1} E_Z &= X_1 (X_1' X_1)^{-1} X_1' Z (Z' Z)^{-1} Z' \\ &= 0. \end{aligned} \quad (39)$$

Substituting Z to X_2 in (36), $C_{Y Y, X_1 X_2}$ vanishes letting $C_{Y Y, X_1}^- X_2$ equal to $C_{Y Y, X_1}$. And, by substituting Z into X_2 , (35) is reduced as follows.

$$\begin{aligned} C_{Y Y, X_1}^- Z &= \frac{1}{N} Y' E_{X_1}^- E_Z E_{X_1}^- Y = \frac{1}{N} Y' E_Z Y \\ &= \frac{1}{N} Y' Z (Z' Z)^{-1} Z' Y \\ &= \frac{1}{N} Y' E_{X_1}^- X_2 (X_2' E_{X_1}^- X_2)^{-1} \\ &\quad \times X_2' E_{X_1}^- Y \\ &= (C_{Y X_2} - C_{Y X_1} C_{X_1 X_1}^{-1} C_{X_1 Y}) \\ &\quad \times (C_{X_2 X_2} - C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2})^{-1} \\ &\quad \times (C_{X_2 Y} - C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2}) \quad (40) \end{aligned}$$

From (37) and (39), we get

$$\begin{aligned} C_{Y Y, X_1}^- Z &= \frac{1}{N} Y' E_{X_1}^- E_Z E_{X_1}^- Y \\ &= \frac{1}{N} Y' (I_N - E_{X_1} - E_Z) Y \\ &= C_{Y Y} - C_{Y Y, X_1} - C_{Y Y, X_1 Z}. \quad (41) \end{aligned}$$

Then $C_{Y Y}$ can be decomposed into three components,

$$C_{Y Y} = C_{Y Y, X_1} + C_{Y Y, X_1}^- Z + C_{Y Y, X_1}^- Z^-. \quad (42)$$

In this case, factors extracted from three component matrices are mutually independent.

In case more than three sets of variables are available

Let's consider the case where the fourth set of variables X_3 is available in addition to three variables sets (Y, X_1, X_2) . By constructing the projection E_{X_3} and premultiplying (30) and (31) by E_{X_3} , we can decompose the total covariance matrix into eight parts. However, it takes very tedious complications to get a complete solution.

In order to avoid these complications, we found that the group factor analytic method was very useful to get a solution of (42) with more than three sets of variables.

For simplicity, we treat the case in which three variables sets Y, X_1 and X_2 are available. Of course this technique can easily be extended to more than three sets. It should be noted, however, that this technique cannot be applied to the case of (38).

Let C be a supper matrix defined as

$$C = \begin{pmatrix} C_{X_1 X_1} & C_{X_1 X_2} & C_{X_1 Y} \\ C_{X_2 X_1} & C_{X_2 X_2} & C_{X_2 Y} \\ C_{Y X_1} & C_{Y X_2} & C_{Y Y} \end{pmatrix} \quad (43)$$

Then we describe the computational procedure in the following manner.

First, apply the successive factor extraction method to $C_{X_1 X_1}$ until all the factors are extracted. When the residual matrix of $C_{X_1 X_1}$ is obtained, the components of X_2 and Y predictable from the extracted factors should be eliminated. This technique correspond to the rank reduction method of the group factor analysis. When all the factors of $C_{X_1 X_1}$ are extracted perfectly, the residual covariance matrix of $C_{X_2 X_2}$ and $C_{Y Y}$ become equal to $C_{X_2 X_2, X_1}^-$ and $C_{Y Y, X_1}^-$ respectively. By setting $\mathbf{C}_{X_2 X_2}$ and $\mathbf{C}_{Y Y}$ to be the residual matrices, we have

$$\begin{aligned} \mathbf{C}_{X_2 X_2} &= C_{X_2 X_2} - C_{X_2 X_1} C_{X_1 X_1}^{-1} C_{X_1 X_2} \\ &= C_{X_2 X_2, X_1}^- \end{aligned} \quad (44)$$

and

$$\begin{aligned}\mathbf{C}_{YY} &= C_{YY} - C_{YX_1}C_{X_1X_1}^{-1}C_{X_1Y} \\ &= C_{YY.X_1}^{-}.\end{aligned}\quad (45)$$

The proof will be summarized as follows.

Let F be the factor score matrix of X_1 , then $F = X_1 T_1 J_1^{-1}$ where J_1^2 and T_1 denote the diagonal matrix with the eigenvalues of $C_{X_1X_1}$ and the square orthogonal matrix with eigenvectors of $C_{X_1X_1}$ respectively.

We show that the projection of F is identical to that of X_1 below.

$$\begin{aligned}E_F &= F(F'F)^{-1}F' = X_1 T_1 J_1 \\ &\quad \times (J_1^{-1} T_1' X_1' X_1 T_1 J_1^{-1})^{-1} J_1^{-1} T_1' X_1' \\ &= X_1 (X_1' X_1)^{-1} X_1' = E_{X_1}\end{aligned}\quad (46)$$

From this relation, the residual covariance matrix of $\mathbf{C}_{X_1X_1}$ after eliminating the effect F becomes equal to $C_{X_2X_2.X_1}^{-}$, since

$$\begin{aligned}C_{X_2X_2.F}^{-} &= \frac{1}{N} X_2' E_F^{-} X_2 = \frac{1}{N} X_2' E_{X_1}^{-} X_2 \\ &= \frac{1}{N} X_2' X_2 - \frac{1}{N} X_2' E_{X_1} X_2 \\ &= C_{X_2X_2} - C_{X_2X_1} C_{X_1X_1}^{-1} C_{X_1X_2} \\ &= C_{X_2X_2.X_1}^{-}.\end{aligned}\quad (47)$$

As for $\mathbf{C}_{YY}$, the same argument holds.

Secondly, we consider to eliminate the predictable parts of $E_{X_1}^{-} Y$ from the factor scores of $E_{X_1}^{-} X_2$.

Setting Z to $E_{X_1}^{-} X_2$, then the projection of the factor score matrix $\mathbf{F}$ of Z becomes identical to that of Z under the same kinds of derivations in (46). This is shown as

$$\begin{aligned}E_F &= E_Z = Z(Z'Z)^{-1}Z' \\ &= E_{X_1}^{-} X_2 (X_2' E_{X_1}^{-} X_2)^{-1} X_2' E_{X_1}^{-}.\end{aligned}\quad (48)$$

From this relation, the residual covariance matrix of $\mathbf{C}_{YY}$ after eliminating the effect of $\mathbf{F}$ becomes

$$\begin{aligned}\mathcal{C}_{YY} &= \frac{1}{N} Y' E_{X_1}^{-} E_F E_{X_1}^{-} Y \\ &= \frac{1}{N} Y' E_{X_1}^{-} E_Z E_{X_1}^{-} Y\end{aligned}$$

$$\begin{aligned}&= \mathbf{C}_{YY} - \frac{1}{N} Y' (I_N - E_{X_1}) X_2 \{ X_2' (I_N \\ &\quad - E_{X_1}) X_2 \}^{-1} X_2' (I_N - E_{X_1}) Y.\end{aligned}\quad (49)$$

This expression becomes identical to $C_{YY.X_1}^{-}$ as shown in (40). Then we have

$$\begin{aligned}\mathbf{C}_{YY} - \mathcal{C}_{YY} &= C_{YY.X_1}^{-} - C_{YY.X_1}^{-} \\ &= C_{YY.X_1}^{-}.\end{aligned}\quad (50)$$

From (44), (45) and (50), we can show that C_{YY} can be decomposed into three parts shown in (42) and C_{XX} can be decomposed into two parts shown in (9) simultaneously.

By applying this technique to more than three sets of variables, we can get the factor solutions related to these sets very easily.

NUMERICAL EXAMPLES

The following examples have already been published elsewhere (cf. Yanai, 1969; Yanai, Hidano, Nushi, Shigematsu & Takane, 1971). However, in order to show the usefulness of our method, we will try to make its brief explanation. The detailed results will be found in the reference above.

Ex. 1. *Construction of a psychological test scale free from social desirability effect.*³

Edwards (1957) assumed that responses to psychological test scale items are liable to be distorted by social desirability (we hereafter call S.D.). Under the assumption above, we try to apply our method as shown in (9) to the construction of a scale free from this effect. For this purpose, we collected the data of responses to all the items by the five rating scale asking the subjects how much each item sounds to be desirable. Based on the data, we define a social desirable score D_i for each item.

³ This is a part of the results concerning the construction of Todai vocational aptitude test. This work is done by Prof. T. Hidano, H. Yanai, T. Nushi, K. Shigematsu and Y. Takane.

By making use of the score, we propose S.D. score for each subject as follows.

Let y_{ik} be the response score of item i by the subject k , we defined S.D. score such that

$$x_k = \frac{\sum_i (D_i - \bar{D})(y_{ik} - \bar{y}_i)}{\sigma_D \sigma y_k} \quad (50')$$

where σ_D and σy_k denote the standard deviations of D_i and y_k respectively. This means that the S.D. score for each subject can be defined as the correlation coefficient between the S.D. score and the response to each item.

Assuming (50) to be an external criterion, we try to apply the principle shown in (9). As the number of the external criteria is one in this case, (9) is replaced by the following formula which is represented in correlational form.

$$R_{YY} = R_{Yx}R'_{Yx} + (R_{YY} - R_{Yx}R'_{Yx}) \quad (51)$$

where R_{YY} denotes the correlation matrix of the items which are assumed beforehand to construct a scale and R_{Yx} denotes a vector, elements of which are correlation coefficients between the S.D. score and items.

One factor extracted from the first term is dependent of the S.D. effect and all the factors extracted from the second term are independent of it.

As an example, we present the results of two scales in Fig. 2 based on the factor loadings dependent of S.D. and the principal factor loadings independent of it together with the principal factor loadings extracted from R_{YY} by means of conventional method.

The results suggest that the items (5), (8), (11), (14), (15), and (18) seem to

(1) Biology-Chemistry Scale

- 1) do the experiment of chemistry
- 2) collect the specimen of rare animals
- 3) investigate the structure of plants
- 4) invent a newchemical fiber
- 5) think of the scientific method of medical diagnosis
- 6) observe the growth of life of the microorganism through microscope
- 7) observe the growing process of the creature
- 8) study to invent the effective medicine for the cancer

(2) Social Extrovert Scale

- 9) be often at the disadvantage because of the retiring disposition
- 10) wish that I were not so shy
- 11) don't speak so much at the meeting and the like
- 12) soon become friendly to a new acquaintance
- 13) rarely speak to a stranger of my own accord
- 14) suffer troubles to find out the appropriate topic for conversation
- 15) cannot express well what I have in mind
- 16) have a lot of friends of the other sex
- 17) think that I am a joy and jovial person
- 18) can associate widely with people of various standpoint

FIG. 2. Factor Loadings of the items.

have high S.D. factor loadings. This seems to mean that the responses to the above items are proved to be easily affected by S.D. effect which is assumed to be independent of the components that the scale items intend to measure. In other words, it seems to be reasonable that the obtained factors through our method are valid *psychometrically* and psychologically interpretable in that the contaminated S.D. effects are taken off.

Ex. 2. *Factor analysis of vocational aptitude test*

We applied three way analysis of variance technique as shown in (27) to the data of the General Aptitude Test Battery (GATB) consisting of 11 scales. The subjects are 950 Japanese students who are classified with respect to sex, developmental (high school and middle school) and regional (rural and urban) difference.

We analyzed eight component matrices separately by the principal component method without estimating communalities. As a result, every component matrix produced only one factor corresponding to the degree of freedom for it. Then, we have seven component factors. On the other hand, the remaining residual matrix R_{ABCe} produced eleven factors. In Fig. 3,

FIG. 3. Eigenvalues obtained by means of two analysis methods

we show the eigenvalues in the descending order of their magnitudes together with eleven factors which were the result by applying the conventional factor analysis to the original correlation matrix R . When we compare the behavior of eigenvalues by our method to the one by the conventional method, it is found that the first eigenvalue of the former is smaller than the latter's one. However, the remaining factors are not so different with respect to their magnitudes. The smaller eigenvalue of the principal factor by our method seem to be dependent of developmental difference, since the eigenvalue corresponding to the component matrix R_B is found to be larger compared to the eigenvalues corresponding to the other component matrices. This means that the factor related to the developmental difference is supposed to be eliminated from the principal factor⁴ by conventional factor analysis based on correlation matrix R .

Through this experiment, we have attained to the idea that the main factor through usual factor analysis seems to be too large and that it has a possibility to be overestimated by practical factor analysis.

REFERENCE

- BURT, S. C. 1966 The appropriate uses of factor analysis and analysis of variance. Cattell, R.C. (Eds.). *Handbook of multivariate experimental psychology* Chicago: Rand McNally, Pp. 267-287.
- CRONBACH, L. J. 1951 Coefficient alpha and internal structure of tests. *Psychometrika*, **16**, 297-334.
- CRONBACH, L. J. & GLESER, G. L. 1965 Generalizability of scores influenced by multiple sources of variance. *Psychometrika*, **3**, 395-420.
- EDWARDS, A. L. 1957 *The social desirability variable in personality assessment research*. New York: The Deyden Press.

⁴ In this study, the principal factor seems to correspond to the general aptitude ability of this test.

- GOLLOB, H. F. 1969 A statistical model which combines features of factor analysis and analysis of variance technique. *Psychometrika* **34**, 73-115.
- GULLIKSEN, H. 1950 *Theory of mental test*. New York: Wiley
- JONES, L. V. 1966 Analysis of variance in its multivariate development. Cattell, R.C. (Eds.). *Handbook of multivariate experimental psychology*. Chicago: Rand McNally, Pp. 244-265.
- TUKEY, J. W. 1962 The future of data analysis. *Annals of mathematical statistics*, **33**, 1-67.
- YANAI, H. 1969 Factor analysis of vocational aptitude test: Application of variance analysis technique to factor analysis. *Jap. J. educ. Psychol.* **17**, 1-12 (in Japanese).
- YANAI, H., HIDANO, T., NUSHI, A., SHIGEMASU, K., & TAKANE, Y. 1971 A methodological study of the test scale construction. *Jap. J. educ. Psychol.* **19**, 33-47 (in Japanese).

(Received Nov., 24, 1970)