R E P O R T   R E S U M E S

ED 010 599   24

A STUDY OF AN EXPLORATORY TECHNIQUE USED IN EDUCATIONAL RESEARCH.

BY- DARROCH, JOHN N.
MICHIGAN UNIV., ANN ARBOR

REPORT NUMBER CRP-S-431   PUB DATE   66
REPORT NUMBER BR-5-8372
EDRS PRICE MF-$0.09 HC-$1.08   27P.

DESCRIPTORS- *EDUCATIONAL RESEARCH, MATHEMATICAL CONCEPTS, *MATHEMATICAL APPLICATIONS, *MATHEMATICAL MODELS, *FACTOR ANALYSIS, *CALCULATION, MEASUREMENT TECHNIQUES, ANN ARBOR, MICHIGAN

TO DEVELOP A BETTER MATHEMATICAL BASIS FOR FACTOR ANALYSIS, A MATHEMATICAL APPROACH WAS FORMULATED FOR FIXING COMMUNALITIES AND THEIR COMPLEMENTS, UNIQUENESSES, AS WELL. WHILE THE INVESTIGATOR WAS UNABLE TO PROVE THAT MINIMUM-TRACE CRITERIA FIX COMMUNALITIES UNIQUELY, A SET OF IDENTITIES AND INEQUALITIES WAS LOCATED AMONG THE COMMUNALITIES AND, EQUIVALENTLY, AMONG THE UNIQUENESSES. PROOF AND ADDITIONAL RESULTS CONCERNING THESE IDENTITIES AND INEQUALITIES APPEARED IN AN APPENDIX. THESE IDENTITIES AND INEQUALITIES WERE OF DIRECT SIGNIFICANCE SINCE THEY AFFORDED A METHOD OF OBTAINING AN APPROXIMATE SOLUTION FOR THE MINIMUM-TRACE COMMUNALITIES. NO TIME WAS AVAILABLE DURING THE PERIOD OF THE CONTRACT TO PERFORM THE NUMERICAL CALCULATIONS NECESSARY TO VERIFY THE FINDINGS. (GD)

5-8372

A Study of an Exploratory Technique Used in Educational Research

Cooperative Research Project No. S-431

John N. Darroch

University of Michigan
Ann Arbor, Michigan

1966

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education and Welfare.

Table of Contents

Page	
3	Problem
4	Objectives
4	Related research
5	Findings
8	References
8	Conclusions

Appendix A: Proposal

Appendix B: Some further inequalities and an identity in factor analysis

PROBLEM.

Let $\underset{\sim}{x} = [\, x_1 x_2 \ldots x_p \,]'$ be a vector of standardised (mean zero, variance one) random variables. The factor analysis model for $\underset{\sim}{x}$ is

$$\underset{\sim}{x} = \underset{\sim}{A}\underset{\sim}{f} + \underset{\sim}{z}$$

where $\underset{\sim}{f} = [\, f_1 f_2 \ldots f_q \,]'$, $q < p$, $\underset{\sim}{z} = [\, z_1 z_2 \ldots z_p \,]'$, $\underset{\sim}{A}$ is $p \times q$, $E(\underset{\sim}{f}) = \underset{\sim}{0}$, $E[\underset{\sim}{f}\underset{\sim}{f}'] = I_{q \times q}$, $E[\underset{\sim}{z}\underset{\sim}{z}'] = \Delta$, diagonal. Thus, if $\underset{\sim}{\Sigma}$ denotes the covariance matrix of $\underset{\sim}{x}$,

$$\underset{\sim}{\Sigma} = \underset{\sim}{A}\underset{\sim}{A}' + \underset{\sim}{\Delta}.$$

The communalities $h_1^2, h_2^2, \ldots, h_p^2$ of $x_1, x_2, \ldots, x_p$ are the diagonal elements of $\underset{\sim}{A}\underset{\sim}{A}'$ or, in other words, the fractions of the variances of $x_1, x_2, \ldots, x_p$ attributable to the common factors $f_1, f_2, \ldots, f_q$. The uniquenesses $\delta_1^2, \delta_2^2, \ldots, \delta_p^2$ of $x_1, x_2, \ldots, x_p$ are the diagonal elements of the diagonal matrix $\underset{\sim}{\Delta}$ or, in other words, the fractions of the variances attributable to the "specific variables" $z_1, z_2, \ldots, z_p$.

The problem concerns the choice of the communalities when $\underset{\sim}{\Sigma}$ is given.

OBJECTIVES.

The general objective was to study the consequences of the minimum-trace criterion for communalities. This says: choose $h_1^2, h_2^2, \ldots, h_p^2$ such that their sum is minimised.

The particular objectives were as follows.

(i) To prove or disprove that this criterion fixes the communalities uniquely.

(ii) To investigate the relationship, if any, between the minimum-trace criterion and the minimum-rank criterion which says: choose the communalities so that q is as small as possible.

(iii) To solve for the minimum-trace communalities.

RELATED RESEARCH.

Darroch (1965) gives inequalities which the communalities must satisfy regardless of which criterion is used to select particular values. These inequalities are relevant to objectives (i) and (iii).

After writing the proposal it was discovered that Ledermann (1939) had investigated our objective (ii). He constructs a numerical example of a 5×5 covariance matrix for which the minimum rank, q, is

2 but such that the minimum-trace is attained when $q > 2$. This demonstrates that there is no strict relationship between the two criteria and, in particular, that neither implies the other. Ledermann noted that Thompson (1938) was the first person to propose the minimum-trace criterion. Thus, in the proposal for the present research, the description of this criterion as "new" was ill founded.

References

Darroch, J. N., A set of inequalities in factor analysis, Psychometrika, 1965, 30, 449-453.

Ledermann, Walter, On a problem concerning matrices with variable diagonal elements. Proc. Roy. Soc. Edinburgh, 1939, 60, 1-17.

Thompson, G. H., Maximising the specific factors in the analysis of ability, Brit. Journ. Educ. Psych., 8, 255-264.

FINDINGS.

Much effort was spent on objective (i) but, apart from the unimportant case when $p = 2$, we were unable to prove that the minimum-trace criterion fixes the communalities uniquely.

As pointed out in the previous section, objective (ii) was effectively answered by Ledermann twenty-seven years ago.

The principal findings concern a set of identities and inequalities among the communalities or, equivalently, among the uniquenesses $\delta_1^2, \delta_2^2, \ldots, \delta_p^2$. These are p identities $E_1, E_2, \ldots, E_p$ of which the first is

$$E_1: 1 - \rho_1^2 = \delta_1^2 + \sum_{r=0}^{\infty} \underset{\sim}{\sigma}_1' (\underset{\sim}{\Sigma}_{11}^{-1} \underset{\sim}{\Delta}_{11})^r \underset{\sim}{\Sigma}_{11}^{-1} \underset{\sim}{\sigma}_1$$

where

$$\underset{\sim}{\Sigma} = \begin{bmatrix} 1 & \underset{\sim}{\sigma}_1' \\ \underset{\sim}{\sigma}_1 & \underset{\sim}{\Sigma}_{11} \end{bmatrix}, \quad \underset{\sim}{\Delta} = \begin{bmatrix} \delta_1^2 & \underset{\sim}{0}' \\ \underset{\sim}{0} & \underset{\sim}{\Delta}_{11} \end{bmatrix}.$$

E_1 implies the inequalities

$$I_1(n): 1 - \rho_1^2 \geq \delta_1^2 + \sum_{r=1}^{n} \underset{\sim}{\sigma}_1' (\underset{\sim}{\Sigma}_{11}^{-1} \underset{\sim}{\Delta}_{11})^r \underset{\sim}{\Sigma}_{11}^{-1} \underset{\sim}{\sigma}, \quad n = 1, 2, 3, \ldots .$$

The proofs and additional results concerning these identities and inequalities have been written up in a form suitable for publication in Appendix B .

The identities and inequalities are of general significance for factor analysis since they show exactly what lies behind the inequalities

$$1 - \rho_i^2 \geq \delta_i^2 , \quad i = 1, 2, \ldots, p ,$$

first proved in 1935 by Roff. They are of direct significance for objective (iii) of this project since they afford a method of obtaining an approximate solution for the minimum-trace communalities, as follows. Let $R(n)$ denote the region of points $(\delta_1^2, \delta_2^2, \ldots, \delta_p^2)$ satisfying both

$$\delta_i^2 \geq 0 , \quad i = 1, 2, \ldots, p ,$$

and the inequalities

$$I_i(n) , \quad i = 1, 2, \ldots, p .$$

The maximation of $\delta_1^2 + \delta_2^2 + \ldots + \delta_p^2$ (equivalently the minimisation of $h_1^2 + h_2^2 + \ldots + h_p^2$) in $R(n)$ is a problem in nonlinear programming (linear if $n = 1$) and can be investigated by standard methods. See for instance Graves and Wolfe (1963). It seems clear that the accuracy of this solution increases with n and that, as $n \to \infty$, the approximate solution converges to the exact solution. However no time was available during the period of the contract to perform the numerical calculations which are necessary to verify the above statements.

References

Graves, Robert L. and Wolfe, Philip, Recent advances in mathematical programming, McGraw-Hill, 1963.

Roff, M., Some properties of the communality in multiple factor theory, Psychometrika, 1936, 1, 1-6.

CONCLUSIONS.

More work is necessary to answer the questions: are the minimum-trace communalities unique and how can they be solved most efficiently? In this work the identities and inequalities given in Appendix B should prove invaluable.

Appendix A
Proposal S-412-65
Project S-431

Project Title: A Study of an Exploratory Technique Used in Educational Research

Submitted by; The University of Michigan
Ann Arbor, Michigan

Initiated by: John N. Darroch
John N. Darroch
Associate Professor

Sponsored by: ______________________
George E. Hay
Chairman of the Department of Mathematics

Transmitted by: ______________________
Robert C. Fitzpatrick
Assistant Director
Office of Research Administration

Date Transmitted: February 17, 1965 (Revised September 23, 1965)

Fiscal Officer: W. K. Pierpont, Vice-President for Business & Finance

Address: 2008 Administration Building, Ann Arbor, Michigan 48105

Phone: Area Code 313, 764-7272

1. ABSTRACT

(a) The objectives are to study the consequences of a new criterion for communalities. In particular (i) to discover whether the criterion defines communalities uniquely in all cases; (ii) to investigate the connection with the usual, but unsatisfactory, minimum-rank definition; (iii) to obtain an accurate and speedy method for solving for the communalities numerically.

(b) The procedure for objectives (i) and (ii) will be mathematical and, for (iii) will be partly mathematical and partly computational.

2. PROBLEM

Factor Analysis is an exploratory technique used in Educational Research and in many other important fields. A factor analysis of a correlation matrix of a set of variables $x_1, x_2, \ldots, x_p$ attempts to explain the correlation of the variables by their dependence on a set of common factors $f_1, f_2, \ldots f_q$. It is similar to a regression analysis but with the important difference that the "independent variables" $f_1, f_2, \ldots, f_q$ are not observed and are sought for beneath the surface of the observations.

Thus, in a recent paper in the Journal of Educational Psychology (1963), Wallen, Travers, Teid and Wodtke are able to represent twenty variables describing teacher behavior in terms of five underlying factors which they label: cold, controlling versus warm, permissive; vigorous, dynamic versus dull, quiet; insecure, anxious versus confident; spends much time alone versus little time alone; much academic activity versus little academic emphasis.

Factor Analysis is more ambitious than most other types of statistical analysis and, partly because of this, there are still some basic ambiguities remaining despite all of the attention it has received since its inception by Charles Spearman i 1904. One of these ambiguities concerns the determination of the communalities $h_1^2, h_2^2, \ldots h_p^2$. The communality h_i^2 of the variable x_i is the fraction of its variance attributable to the common factors $f_1, f_2, \ldots, f_q$. No completely satisfactory criterion for determining the communalities has yet been offered and this is the basic problem which I wish to tackle. It might seem surprising that factor analysis has been used so much when there is this fundamental weakness in the theory. However, there is much numerical evidence to suggest that the conclusions of a factor analysis are not greatly altered by changing the communalities and it is this fact which permits many users of factor analysis to forget that the theory is intrinsically unsound.

I wish to propose a new criterion for communalities which I believe will provide an acceptable solution to the problem.

3. LITERATURE

Factor Analysis is used in Educational Research so extensively that it would be impertinent to try and summarize the relevant literature. Practically every issue of the Journal of Educational Psychology and many issues of "Educational and Psychological Measurement" and the Journal of Educational Research report at least one piece of research in which Factor Analysis was a vital tool.

Turning to the general theory, the guiding principle is that the factor model should be parsimonious in the sense that q should be as small as possible. (The question of whether Nature operates in this parsimonious way is an interesting one but one which cannot usually be answered.) If Γ denotes the modified correlation matrix in which the diagonal elements of one are replaced by the communalities, then q is the rank of Γ. Therefore, by appealing to the "principle of parsimony," we can say that the rank of Γ should be minimized with respect to $h_1^2, h_2^2, \ldots, h_p^2$ subject to the condition that Γ is kept non-negative definite (i.e. representable as a covariance matrix). In some cases this minimum rank condition fixes $h_1^2, h_2^2, \ldots, h_p^2$ uniquely but in the majority of cases it <u>does not</u>; Anderson and Rubin (1956) is the most important reference here. Moreover, even when the condition of minimum rank does lead to unique values of $h_1^2, h_2^2, \ldots, h_p^2$, there is no direct method of solving for these values.

Because of these difficulties, other criteria for communalities have been offered. Guttman (1956) proposed that the "best possible" estimate of h_i^2 is ρ_i^2 where ρ_i is the multiple correlation coefficient of x_i with the remaining p-1 variables. However, this leads to a Γ which is illegitimate in the sense of not being non-negative definite. To meet this drawback Joreskog (1963) has proposed that $1-h_i^2 = \Theta\ (1-\rho_i^2)$ where Θ is the largest number which leaves Γ non-negative definite. Both of these proposals are manageable computationally and their intuitive content derives from two sources. Firstly, the inequalities $h_i^2 \geq \rho_i^2$ which were first pointed out by Roff (1936). Secondly, Guttman's theorem (1956) that, under very reasonable conditions, h_i can be viewed as the multiple correlation coefficient of x_i with an <u>infinite</u> set of other relevant variables. Some tighter inequalities than the above are presented in the attached note.

Likelihood-ratio hypothesis-testing provides an approximate decomposition of the sample correlation matrix which parallels the exact decomposition of the population correlation matrix into Γ plus the diagonal matrix of "uniquenesses" $1-h_i^2$. But the fact that the likelihood-ratio decomposition is approximate and assumes a normal distribution for the observations means that it is not strictly relevant to the present discussion.

A closely related alternative to Factor Analysis is Image Analysis proposed by Guttman (1953). Harris (1962) has shown that Image Analysis is closely related to a factor analysis which uses Guttman's best possible communalities. In the same paper Harris also relates Guttman's work to some proposals of Rao (1955) for deriving communalities iteratively once q has been decided upon.

A good overall discussion of the communality problem is to be found in Chapter 5 of the book by Harman (1960).

References

Anderson, T.W. and Rubin, H. (1956). Statistical inference in factor analysis. Proc. Third Berkeley Symposium 5, 111-150.

Darroch, J.N. A set of inequalities in factor analysis. (See attached notes.)

Guttman, L. (1953). Image theory for the structure of quantitative variates. Psychometrika, 18, 277-96.

Guttman, L. (1956) "Best possible" systematic estimates of communalities. Psychometrika, 21, 273-85.

Harman, H.H. (1960) Modern factor analysis. University of Chicago Press.

Harris, C. (1962) Some Rao-Guttman relationships. Psychometrika, 27, 247-263.

Joreskog, K.G. (1963) Statistical estimation in factor analysis: a new technique and its foundation. Stockholm, Almquist and Wicksell.

Rao, C.R. (1955) Estimation and tests of significance in factor analysis. Psychometrika, 20, 93-111.

Roff, M. (1935) Some properties of the communality in multiple factor theory. Psychometrika, 1, 1-6.

4. OBJECTIVES

I propose the following criterion for communalities: that they be chosen to minimize their sum $h_1^2 + h_2^2 + \ldots + h_p^2 = \text{trace } \Gamma$, subject to Γ being positive semi-definite. This definition can be viewed as an interpretation of the principle of parsimony and it dovetails closely with the principal components analysis of Γ; a paper on this aspect of the minimum-trace criterion is under preparation.

My future objectives are to investigate the consequences of this criterion under the following heads.

(i) Uniqueness of $h_1^2, h_2^2, \ldots, h_p^2$. I hope to prove that the communalities are fixed uniquely.

(ii) Relationship between minimum rank and minimum trace. Preliminary investigation indicates a close relationship and it may even be true that minimum trace implies minimum rank.

(iii) Methods of solution for $h_1^2, h_1^2, \ldots, h_p^2$.

5. PROCEDURES

(a) The essential problem involved in 4 (iii) is to provide a good approximation to the smallest root of a positive definite matrix. I have derived an upper bound for this root which is intimately related to the inequalities proved in the attached note; also a lower bound. Apart from possible theoretical refinement of these bounds the main task will be to investigate numerically the accuracy of the communality solutions derived from them. This numerical work will require some time on The University of Michigan IBM 7090 computer.

(b), (c), (d). Not applicable.

(e) Not possible to give any reliable estimates.

6. PERSONNEL

The principal investigator meets regularly with the mathematical psychologists at The University of Michigan and shall address their weekly seminar in the spring semester. Professor Paul Dwyer of the Department of Mathematics and the Statistics Laboratory stated that he wishes to cooperate on the computational aspects of this project. Among other things Professor Dwyer is a foremost worker in computational statistics and was responsible for some of the early mathematical development of Factor Analysis.

Biographical data concerning the principal investigator follows.

BIOGRAPHICAL DATA

DARROCH, JOHN NEWTON — Visiting Lecturer

Education:

B. A., Peterhouse, Cambridge	1952
Diploma in Statistics, Peterhouse, Cambridge	1953
Ph.D., Capetown University	1960

Employment: Academic

Cambridge University, Cambridge, England,	Graduate Student	1952-53
R. A. F. Technical College, Henlow, Beds.,	Flying Officer - Education Branch	1953-55
Capetown University, Capetown, S. Africa	Lecturer in Math.	1955-58
Manchester University, Manchester, England	Lecturer in Math. Statistics	1959-62
Adelaide University, Adelaide, South Australia,	Senior Lecturer in Math. Statistics	1962-64

Bibliography

The multiple-recapture census I estimation of a closed population, Biometrika, 45, 343-59 (1958).

The multiple-recapture census II estimation when there is immigration or death, Biometrika, 46, 336-51 (1959).

The two-sample capture-recapture census when sampling and tagging are stratified, Biometrika, 48, 541-60 (1961).

Interactions in multi-factor contingency tables, J. R. S. S. (B), 24, 251-263 (1962).

On testing more than one hypothesis, with S. D. Silvey, Ann. Math. Statist. 34, 555-67 (1963).

On the traffic light queue, Ann. Math. Statist. 35, 380-388 (1964).

On the distribution of the number of successes in independent trials, Ann. Math. Statist., 35, 1317-1321 (1964).

Queues for a vehicle-actuated traffic light, with A. F. Newell and R. W. J. Morris, J. Operations Research 12, 882-895 (1964).

On Quasi-Stationary Distributions in Absorbing Discrete-Time Finite Markov Chains, with E. Seneta. Accepted for publication in J. Applied Probability (1965) (20 typescript pages).

A Set of Inequalities in Factor Analysis, submitted to Psychometrika (7 typescript pages).

7. FACILITIES

These include the excellent Mathematics and Statistics Library at The University of Michigan, as well as the services of the Statistical Laboratory and the Computing Center which contains an IBM 7090 computer.

8. OTHER INFORMATION

(a) No support is available for this project from sources other than the transmitting institution.

(b) This proposal has not been submitted to any other agency or organization.

(c) This is not a proposed extension of, or addition to, a previous or current project sponsored by the Office of Education or any other group or agency.

(d) This proposal, or one similar to it, has not been previously submitted to the Office of Education.

SOME FURTHER INEQUALITIES AND AN IDENTITY IN FACTOR ANALYSIS

John N. Darroch

1. Introduction and Summary

As in [1] let Σ denote the (non-singular) covariance matrix of $x = [x_1\, x_2 \cdots x_p]'$ where x_i has mean zero and variance one, $1 \leq i \leq p$. The first basic requirement of a factor-analysis model for x is that x can be expressed as

$$x = y + z,$$

$y = [y_1\, y_2 \cdots y_p]'$, $z = [z_1\, z_2 \cdots z_p]'$, where

$$E(y) = 0, \quad E(z) = 0$$

and

(1) $\quad E[y z'] = 0, \quad E[z z'] = \Delta,$ diagonal.

Implicit in (1) is the condition that the matrix $\Sigma - \Delta$ is non-negative definite since it can be written as $E[y y']$.

Write

$$x_1 = [x_2\, x_3 \cdots x_p]', \quad y_1 = [y_2\, y_3 \cdots y_p]', \quad z_1 = [z_2\, z_3 \cdots z_p]',$$

$$\Sigma_{11} = E[x_1 x_1'], \quad \Delta_{11} = E[z_1 z_1'],$$

$$\beta_1 = \Sigma_{11}^{-1} \sigma_1, \quad \rho_1^2 = \sigma_1' \Sigma_{11}^{-1} \sigma_1 = \beta_1' \Sigma_{11} \beta_1.$$

The vector β_1 is the vector of regression coefficients of x_1 on x_1 and ρ_1 is the multiple correlation coefficient of x_1 with x_1.

In [1] we derived inequalities $I_1, I_2, \cdots, I_p$, say, where

$$I_1 : 1 - \rho_1^2 \geq \delta_1^2 + \beta_1' \Delta_{11} \beta_1.$$

δ_1^2 is the (1, 1) element of Δ, the "uniqueness" of x_1. I_1 is an improvement over

$$J_1 : 1 - \rho_1^2 \geq \delta_1^2$$

which was previously derived by Roff [2]. It was shown that the conditions, C_1 say, for equality in I_1 are the same as the conditions for equality in J_1.

In section 2 of this paper the improvement from J_1 to I_1 is carried further, to $I_1(n)$ with conditions $C_1(n)$ for equality. It is shown that

$$C_1(n) \Leftrightarrow C_1, \quad n = 1, 2, \cdots .$$

In section 3 the "ultimate" inequality $I_1(\infty)$, obtained by letting $n \to \infty$ in $I_1(n)$, is considered and the condition for equality is shown to be $C_1(\infty)$: y_1 is a linear combination of $y_2, y_3, \cdots, y_p$. This condition will be recognized as the second basic requirement of a factor-analysis model. [The variables $y_1, y_2, \cdots, y_p$ are viewed as linear combinations of at most $p - 1$ common factors all of which must appear with non-zero coefficients in at least one of $y_2, \cdots, y_p$: otherwise they would not be common.] Thus $I_1(\infty)$ really states an identity, $E_1(\infty)$ say, and it is this identity which is the most fundamental result given here, especially since all of the inequalities $J_1, I_1, I_1(1), \cdots, I_1(n), \cdots,$ are derivable from it.

2. The Inequality $I_1(n)$ and Conditions for Equality.

Define

$$\alpha_1(n) = [I + (\Sigma_{11}^{-1}\Delta_{11}) + \cdots + (\Sigma_{11}^{-1}\Delta_{11})^n]\beta_1 .$$

Then

$$(2)\quad [\Sigma - \Delta]\begin{bmatrix} 1 \\ \cdots\cdots \\ -\alpha_1(n) \end{bmatrix} = \begin{bmatrix} 1 - \rho_1^2 - \delta_1^2 - \beta_1'\Delta_{11}[I + (\Sigma_{11}^{-1}\Delta_{11}) + \cdots + (\Sigma_{11}^{-1}\Delta_{11})^{n-1}]\beta_1 \\ \Delta_{11}(\Sigma_{11}^{-1}\Delta_{11})^n\beta_1 \end{bmatrix}$$

and

$$(3)\quad [1 \,\vdots\, -\alpha_1(n)][\Sigma - \Delta]\begin{bmatrix} 1 \\ \cdots\cdots \\ -\alpha_1(n) \end{bmatrix} = 1 - \rho_1^2 - \delta_1^2 - \beta_1\Delta_{11}[I + (\Sigma_{11}^{-1}\Delta_{11}) + \cdots + (\Sigma_{11}^{-1}\Delta_{11})^{2n}]\beta_1 .$$

Because $\Sigma - \Delta$ is non-negative definite, it follows that the right side of (3) is non-negative. This gives us the inequality $I_1(2n)$ where $I_1(n)$ is defined as

$$I_1(n) : 1 - \rho_1^2 \geq \delta_1^2 + \beta_1'\Delta_{11}[I + (\Sigma_{11}^{-1}\Delta_{11}) + \cdots + (\Sigma_{11}^{-1}\Delta_{11})^n]\beta_1 .$$

The matrix Σ_{11}^{-1} is positive definite and therefore

$$\beta_1'\Delta_{11}(\Sigma_{11}^{-1}\Delta_{11})^r\beta_1 \geq 0 .$$

It follows that

$$I_1(n) \Rightarrow I_1(n-1) \Rightarrow \cdots \Rightarrow I_1(0) \Leftrightarrow I_1 ,$$

and that

$$(4)\quad C_1 \Leftrightarrow C_1(0) \Rightarrow C_1(1) \Rightarrow \cdots \Rightarrow C_1(n),$$

where $C_1(n)$ denotes the conditions for equality in $I_1(n)$, $n = 0, 1, 2, \ldots$.

Now note that, if Γ is a non-negative definite matrix

$$\gamma'\Gamma\gamma = 0 \quad \Leftrightarrow \quad \Gamma\gamma = 0 .$$

Equations (1) and (2) therefore show that

(5) $\quad C_1(2n) \Rightarrow C_1(n-1), \qquad n = 1, 2, \cdots,$

and (4) and (5) yield

$$C_1(n) \Rightarrow C_1, \qquad n = 1, 2, \cdots.$$

3. The Inequality $I_1(\infty)$ and the Condition for Equality.

On letting n approach ∞ in $I_1(n)$ we obtain

$$I_1(\infty): 1 - \rho_1^2 \geq \delta_1^2 + \sum_{r=0}^{\infty} \beta_1' \Delta_{11} (\Sigma_{11}^{-1} \Delta_{11})^r \beta_1 .$$

THEOREM. The condition $C_1(\infty)$ for equality in $I_1(\infty)$ is

$$C_1(\infty): \Sigma - \Delta \text{ has the same rank as } \Sigma_{11} - \Delta_{11},$$

or, in other words,

$$C_1(\infty): y_1 \text{ is a linear function of } y_2, \cdots, y_p .$$

Before proving this theorem we state and prove two lemmas.

LEMMA 1. Define

$$\Theta = \Sigma_{11}^{-1/2} \Delta_{11} \Sigma_{11}^{-1/2} .$$

Then, as $n \rightarrow \infty$, Θ^n converges to a matrix Φ, say, where $\Theta \Phi = \Phi$.

Proof. Let θ denote an eigenvalue of Θ. Now Θ is non-negative definite and therefore θ is real and $\theta \geq 0$. Also $\Sigma_{11} - \Delta_{11}$ is non-negative definite and therefore so is $I - \Theta$ and consequently $\theta \leq 1$. Because θ is real and $|\theta| \leq 1$ it follows that

$$\Theta^n \longrightarrow \Phi$$

where $\Theta \Phi = \Phi$ since $\Theta \Theta^n = \Theta^{n+1}$.

The main purpose of lemma 2 is to establish that, while $\sum_{r=0}^{n} \Theta^r$ need not converge as $n \longrightarrow \infty$, the vector

$$\underset{\sim}{g}_1(n) = \underset{\sim}{\Sigma}_{11}^{-1/2} [\sum_{r=0}^{n} \underset{\sim}{\Theta}^r] \underset{\sim}{\Sigma}_{11}^{-1/2} \underset{\sim}{g}_1$$

does converge.

LEMMA 2. As $n \longrightarrow \infty$, $\underset{\sim}{g}_1(n) \longrightarrow \underset{\sim}{g}_1$, say, where

$$[\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}] \underset{\sim}{g}_1 = \underset{\sim}{g}_1 .$$

Proof. From (1) we have

$$E[\underset{\sim}{y}_1 \underset{\sim}{y}_1'] = \underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}, \quad E[\underset{\sim}{y}_1 (y_1 \vdots \underset{\sim}{y}_1')] = [\underset{\sim}{g}_1 \vdots \underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}] .$$

Therefore there exists at least one solution $\underset{\sim}{\gamma}_1$ of

$$[\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}] \underset{\sim}{\gamma}_1 = \underset{\sim}{g}_1 .$$

Therefore

$$\begin{aligned} \underset{\sim}{g}_1(n) &= \underset{\sim}{\Sigma}_{11}^{-1/2} [\sum_{r=0}^{n} \underset{\sim}{\Theta}^r] \underset{\sim}{\Sigma}_{11}^{-1/2} [\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}] \underset{\sim}{\gamma}_1 \\ &= \underset{\sim}{\Sigma}_{11}^{-1/2} [\sum_{r=0}^{n} \underset{\sim}{\Theta}^r] [\underset{\sim}{I} - \underset{\sim}{\Theta}] \underset{\sim}{\Sigma}_{11}^{1/2} \underset{\sim}{\gamma}_1 \\ &= \underset{\sim}{\Sigma}_{11}^{-1/2} [\underset{\sim}{I} - \underset{\sim}{\Theta}^{n+1}] \underset{\sim}{\Sigma}_{11}^{1/2} \underset{\sim}{\gamma}_1 \\ &\longrightarrow \underset{\sim}{\Sigma}_{11}^{-1/2} [\underset{\sim}{I} - \underset{\sim}{\Phi}] \underset{\sim}{\Sigma}_{11}^{1/2} \underset{\sim}{\gamma}_1 = \underset{\sim}{g}_1 \text{ say.} \end{aligned}$$

Although $\underset{\sim}{g}_1$ is expressed as a function of the particular solution $\underset{\sim}{\gamma}_1$ it is of course the same for all $\underset{\sim}{\gamma}_1$. Finally

$$\begin{aligned} [\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}] \underset{\sim}{g}_1 &= \underset{\sim}{\Sigma}_{11}^{1/2} [\underset{\sim}{I} - \underset{\sim}{\Theta}] [\underset{\sim}{I} - \underset{\sim}{\Phi}] \underset{\sim}{\Sigma}_{11}^{1/2} \underset{\sim}{\gamma}_1 \\ &= \underset{\sim}{\Sigma}_{11}^{1/2} [\underset{\sim}{I} - \underset{\sim}{\Theta}] \underset{\sim}{\Sigma}_{11}^{1/2} \underset{\sim}{\gamma}_1 \\ &= [\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}] \underset{\sim}{\gamma}_1 = \underset{\sim}{g}_1 \end{aligned}$$

since $[\underset{\sim}{I} - \underset{\sim}{\Theta}] \underset{\sim}{\Phi} = \underset{\sim}{0}$.

[Note that, if $\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}$ is non-singular, then $0 \leq \theta < 1$, $\underset{\sim}{\Phi} = 0$ and $\underset{\sim}{g}_1 = [\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}]^{-1} \underset{\sim}{g}_1$.]

Proof of the theorem. Consider the squared multiple-correlation coefficient, $\rho^2(y_1 \mid \underset{\sim}{y}_1)$ of y_1 with $\underset{\sim}{y}_1$. It satisfies the equation

$$(1-\delta_1^2)\,\rho^2(y_1 \mid \underset{\sim}{y}_1) = \underset{\sim}{\sigma}_1' \underset{\sim}{\gamma}_1$$

where, as in lemma 2, $\underset{\sim}{\gamma}_1$ is any solution of

$$[\underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Delta}_{11}]\,\underset{\sim}{\gamma}_1 = \underset{\sim}{\sigma}_1 .$$

In particular

$$\begin{aligned}(1-\delta_1^2)\,\rho^2(y_1 \mid \underset{\sim}{y}_1) &= \underset{\sim}{\sigma}_1' \underset{\sim}{\alpha}_1 \\ &= \underset{\sim}{\sigma}_1' \underset{\sim}{\Sigma}_{11}^{-1/2} \sum_{r=0}^{\infty} [\underset{\sim}{\Theta}^r \underset{\sim}{\Sigma}_{11}^{-1/2} \underset{\sim}{\sigma}_1] \\ &= \rho_1^2 + \sum_{r=0}^{\infty} \underset{\sim}{\beta}_1' \underset{\sim}{\Delta}_{11} (\underset{\sim}{\Sigma}_{11}^{-1} \underset{\sim}{\Delta}_{11})^r \underset{\sim}{\beta}_1 .\end{aligned}$$

Thus the inequality $I_1(\infty)$ may be expressed as

$$I_1(\infty):\ \rho^2(y_1 \mid \underset{\sim}{y}_1) \leq 1,$$

and there is equality if and only if y_1 is a linear function of $y_2, \cdots, y_p$. Since this condition is satisfied in a factor analysis model we have the identity

$$E_1(\infty): 1-\rho_1^2 = \delta_1^2 + \sum_{r=0}^{\infty} \underset{\sim}{\beta}_1' \underset{\sim}{\Delta}_{11} (\underset{\sim}{\Sigma}_{11}^{-1} \underset{\sim}{\Delta}_{11})^r \underset{\sim}{\beta}_1 .$$

REFERENCES

[1] Darroch, J. N., A set of inequalities in factor analysis. Psychometrika, 1965, 30, 449-453.

[2] Roff, M., Some properties of the communality in multiple factor theory. Psychometrika, 1935, 1, 1-6.