PSYCHOMETRIKA—VOL. 41, NO. 3
SEPTEMBER, 1976

DETERMINING THE NUMBER OF COMPONENTS FROM THE MATRIX OF PARTIAL CORRELATIONS

WAYNE F. VELICER

UNIVERSITY OF RHODE ISLAND

A common problem for both principal component analysis and image component analysis is determining how many components to retain. A number of solutions have been proposed, none of which is totally satisfactory. An alternative solution which employs a matrix of partial correlations is considered. No components are extracted after the average squared partial correlation reaches a minimum. This approach gives an exact stopping point, has a direct operational interpretation, and can be applied to any type of component analysis. The method is most appropriate when component analysis is employed as an alternative to, or a first-stage solution for, factor analysis.

Key words: common variance, factor analysis, image component analysis.

This paper is concerned with the class of methods that involve an eigen decomposition of a covariance matrix. Collectively, these methods are called component analysis in this paper. The nature of the covariance matrix determines the specific type of component analysis. The component analysis of a correlation matrix is called principal component analysis here. We shall also be interested in image component analysis, the component analysis of image covariances. Component analysis has a variety of purposes. It can be used to find a parsimonious description of the total variance of the variables involved; in this case, the procedure proposed in this paper is not applicable. Principal component analysis is frequently used to express the variance shared among variables in a set; that is, it is used as kind of a factor analysis. Horst [1965] and Van de Geer [1971], among others, discuss principal component analysis employed in this manner.

When component analysis is used to describe the common variance, as a kind of factor analysis, the usual problem of determining how many components to retain arises. Various solutions have been proposed, none of which are totally satisfactory. Bartlett [1950, 1951] has proposed a statistical test of significance that could be applied to either principal component analysis or image component analysis. This approach generally results in the retention of an "excessive" number of components, *i.e.*, components that are statis-

Requests for reprints should be sent to Wayne F. Velicer, Department of Psychology, University of Rhode Island, Kingston, R. I. 02881.

tically significant but of no practical importance [Gorsuch, 1973]. This problem is particularly acute if the sample size is large.

Numerous methods have been proposed to determine the number of non-trivial components as opposed to significant but trivial components. Probably the most widely employed rule for principal component analysis is to retain only factors corresponding to eigenvalues greater than unity. Kaiser [1960] has rationalized this rule in a number of ways. Gorsuch [1974] reviews some of the criticisms of this approach. Alternative methods include using the scree test [Cattell, 1966], computing the percentage of variance extracted, or assessing the patterns produced by varying the number of factors extracted. All of these methods have been criticized for being either too subjective or too arbitrary.

This paper will consider an alternative method based on the matrix of partial correlations. The method gives an exact stopping point and has a direct rationale with respect to a traditional criterion for factor analysis. The method applies to any type of component analysis, including both principal component analysis and image component analysis.

Method

For a set of $p + m$ variables, the covariance matrix can be represented as

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}' & C_{22} \end{bmatrix}, \tag{1}$$

where the submatrices C_{11}, C_{12}, and C_{22} have dimensions $p \times p$, $p \times m$, and $m \times m$. If the m variables of the second set are partialed out from the p variables of the first set, the matrix of partial variances and covariances can be represented as

$$C_{11}^* = C_{11} - C_{12}C_{22}^{-1}C_{12}', \tag{2}$$

and the matrix of partial correlations is

$$R_{11}^* = D^{-1/2}(C_{11} - C_{12}C_{22}^{-1}C_{12}')D^{-1/2}, \tag{3}$$

where

$$D = \text{Diag}\,(C_{11} - C_{12}C_{22}^{-1}C_{12}') = \text{Diag}\, C_{11}^*. \tag{4}$$

For component analysis, let the first set of p variables represent the observed variables and the second set of m variables represent the components. Then, if A is the $p \times m$ pattern matrix resulting from component analysis, the partial covariance matrix can be represented as

$$C_{11}^* = C_{11} - AA'. \tag{5}$$

The nature of C_{11} would depend on the type of component analysis employed. For principal component analysis,

$$C_{11} = R, \tag{6}$$

where R is the $p \times p$ correlation matrix; for the image component analysis,

$$C_{11} = S^{-1}RS^{-1}, \tag{7}$$

where

$$S^2 = \text{Diag}^{-1}\,(R^{-1}). \tag{8}$$

In order to determine the number of components to extract, the following summary statistic is proposed:

$$f_m = \sum_{i \neq j}\sum (r_{ij}{}^*)^2/(p(p-1)), \tag{9}$$

where $r_{ij}{}^*$ is the element in row i and column j of the matrix $R_{11}{}^*$ in (3). The value of f_m is the average of the squared partial correlations after the first m components are partialed out. The proposed stopping point is the value of m for which f_m is at a minimum. The value f_m would be calculated for $m = 1$ to $p - 1$; the value of f_m for $m = p$ is indeterminate since the diagonal of $C_{11}{}^*$ consists of zeros. The values of f_m will range between 0 and 1. A second summary statistic, useful for comparative purposes, is

$$f_0 = \sum_{i \neq j}\sum r_{ij}{}^2/(p(p-1)). \tag{10}$$

If $f_1 > f_0$, then no components would be extracted.

In order to better understand why the value of f_m first decreases and then increases, consider the general form of a partial correlation equation

$$r_{ij.y} = \frac{r_{ij} - r_{iy}r_{jy}}{((1 - r_{iy}{}^2)(1 - r_{jy}{}^2))^{1/2}}, \tag{11}$$

where i, j represent any two of the p observed variables and y represents a component. With no loss of generality, we assume r_{ij}, r_{iy}, and r_{jy} to be positive. The value of the partial correlation will decrease as long as the numerator decreases faster than the denominator. The value of the partial correlation will increase whenever the reverse is true; for example, when r_{iy} is large and r_{jy} is small. Such a situation would occur if a component has a high correlation with only one variable and near zero correlations with the others (commonly referred to as a "specific"). This explanation can be extended inductively to the case where more than one component has been extracted by using k-th order partials calculated from $(k - 1)$th order partials. Using this stopping rule, the first m components could be referred to as "common components". We would expect that two or more variables would have "high" loadings on these components. The remaining $m + 1$ to p components may include some that have adequate reliability. However, these will be "unique" components (*i.e.*, having "high" loadings on only one of the original variables). The value m^*, the number of components with eigenvalues greater than unity, can be interpreted as the lower limit of m ($m \leq m^*$).

Application

The proposed stopping rule was applied to six studies [Emmett, 1949; Maxwell, 1961; Bechtoldt, 1961; Harman, 1960; Thurstone, 1951; and Lord, 1956] previously reanalyzed by Jöreskog [1967], and Schönemann and Wang [1972]. The value of m as determined from the partial correlation matrix and value of m determined by previous studies are compared in Table 1. The values of f_m are presented for both principal component analysis and image component analysis. In general, the proposed stopping rule would suggest extracting fewer factors than previous analyses have done. In addition to the data presented here, the author has applied this stopping rule to over fifty other studies. The Thurstone [1951] data is of particular interest since it represents the rare case where the stopping rule suggests retaining different numbers of components for principal component analysis and image component analysis.

Discussion

The proposed stopping rule has been developed within the context of component analysis. Some of the concepts used implicitly in the development and rationale of this method were borrowed from factor analysis, a method which has a similar purpose and typically produces results which are empirically very similar [Velicer, 1974, 1976, in press]. The factor analytic model specifies that a $p \times p$ correlation (or covariance) matrix can be accounted for by m common factors and p unique factors. This can be represented as

$$R = AA' + U^2, \tag{12}$$

where A is a $p \times m$ matrix of weights for the common factors ("common factor pattern") and U is the $p \times p$ diagonal matrix of weights for the unique factors ("unique factor pattern"). For a given value of m, both A and U^2 must be estimated under some criterion of best fit that is typically statistical rather than exact. A number of alternative criterion have been developed for defining what is the best fit. Two of these employ a concept similar to the criterion suggested here as a stopping rule.

One of the traditional rationales for factor analysis is that by extracting factors the correlation between the observed variables should be reduced. Howe [Note 1] presented a more general derivation of Lawley's maximum likelihood factor analysis employing the criterion that the determinent of the matrix of partial correlations should be maximized. The maximum is achieved when all the partial correlations equal zero. For maximum likelihood, f_m would be a nonincreasing function. Under the "identifiability constraint"

Table 1 Total Off-Diagonal Residuals (Six Studies)

Part 1. Principal Component Analysis

Source	Emmett	Maxwell	Bechtoldt S1	Harman	Thurstone	Lord
N	211	810	212	305	213	649
P	9	10	17	8	13	33
1	.067398	.044653	.042518	.245121	.082152	.061518
2	.045705	.034750	.037996	.066445	.056091	.038279
3	.077405	.059541	.035230	.127594	.059116	.019597
4	.124539	.086835	.038367	.204203	.076860	.013693
5	.214674	.130898	.040368	.271829	.079224	.010010
6	.290397	.208927	.040432		.103467	.011929
7	.506459	.301387	.054676		.137009	.014999
8			.069998			.016209
9			.096135			.018996
10			.111482			.022502
11						.025316

Part 2. Image Component Analysis

1	.078516	.044355	.145613	.152352	.083952	.066234
2	.043910	.035493	.041128	.064612	.059903	.034007
3	.065997	.054899	.030948	.096197	.049168	.019196
4	.105201	.079829	.032101	.149035	.053733	.015252
5	.167824	.122178	.036749	.254251	.063112	.009502
6	.266348	.182330	.039659		.080741	.010205
7	.449028	.282713	.048408		.099655	.011293
8			.060248			.012604
9			.071241			.014040
10			.094659			.015707
11						.017555

Part 3. Numbers of Factors

Previous *	2,3,4	3	5	2,3	4	9,10,11
Partial Correlation	2	2	3	2	2-3	5

*Schonemann and Wang, 1972.

that

$$A'U^{-2}A = D^2 = \text{diagonal}, \tag{13}$$

the solution can be expressed as an eigen decomposition of the (unknown) covariance matrix

$$\underline{U^{-1}}RU^{-1}. \tag{14}$$

The function f_m is nonincreasing because the covariance matrix described in (14) changes as a function of the number of factors extracted. Thus, if the stopping rule were applied to maximum likelihood factor analysis, a total $p - 1$ factors would be extracted, since the estimated U^2 will always serve to further decrease f_m compared to f_{m-1}.

A related approach, again assuming the factor analytic model, is Minres [Harman & Jones, 1966]. This approach seeks to minimize a function of the off diagonal residual covariation, and takes the form

$$tr(C_{11}{}^* - D)'(C_{11}{}^* - D), \tag{15}$$

where $C_{11}{}^*$ and D are defined by (2) and (4). This is a nonincreasing function for both factor analysis and component analysis.

The stopping rule employs the factor analytic concept of "common" factors to define how many components to extract. The particular method used to determine which components are common has also been employed with the factor analytic model, but as a means of defining the best solution rather than as a stopping rule. In addition to a clearly defined rationale, the stopping point is exact and can be applied with equal success to any covariance matrix. Component analysis may be employed for many different purposes. It is not suggested that the stopping rule discussed here is appropriate for all of them. However, if component analysis is employed as an alternative to factor analysis or as a first-stage solution for factor analysis, the stopping rule proposed here would seem the most appropriate.

REFERENCE NOTE

1. Howe, W. G. *Some contributions to factor analysis* (Report No. ONRL-1919). Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1955.

REFERENCES

Bartlett, M. S. Tests of significance in factor analysis. *The British Journal of Psychology*, 1950, *3* (*Part II*), 77–85.

Bartlett, M. S. A further note on tests of significance in factor analysis. *The British Journal of Psychology*, 1951, *4* (*Part I*), 1–2.

Bechtoldt, H. P. An empirical study of the factor analysis stability hypothesis. *Psychometrika*, 1961, *26*, 405–432.

Cattell, R. B. The scree test for the number of factors. *Multivariate Behavioral Research*, 1966, *1*, 245–276.

Emmett, W. C. Factor analysis by Lawley's method of maximum likelihood. *British Journal of Psychology, Statistical Section*, 1949, 2, 90–97.

Gorsuch, R. L. Using Bartlett's significance test to determine the number of factors to extract. *Educational and Psychological Measurement*, 1973, *33*, 361–364.

Harman, H. H. *Modern factor analysis*. Chicago: University of Chicago Press, 1960.

Harman, H. H., & Jones, W. H. Factor analysis by minimizing residuals (minres). *Psychometrika*, 1966, *31*, 351–368.

Horst, P. *Factor analysis of data matrices*. New York: Holt, Rinehart, and Winston, 1965.

Jöreskog, K. G. Some contributions to maximum likelihood factor analysis. *Psychometrika*, 1967, *32*, 443–482.

Kaiser, H. F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 1960, *20*, 141–151.

Lord, F. M. A study of speed factors in tests and academic grades. *Psychometrika*, 1956, *21*, 31–50.

Maxwell, E. A. Recent trends in factor analysis. *Journal of the Royal Statistical Society*, Series A, 1961, *124*, 49–59.

Schönemann, P. H., & Wang, M. Some new results on factor indeterminacy. *Psychometrika*, 1972, *37*, 61–91.

Thurstone, L. L. The dimensions of temperament. *Psychometrika*, 1951, *16*, 11–20.

Van de Geer, J. P. *Introduction to multivariate analysis for the social sciences*. San Francisco: Freeman, 1971.

Velicer, W. F. An empirical comparison of the stability of factor analysis, principal component analysis, and image analysis. *Educational and Psychological Measurement*, 1974, *34*, 563–572.

Velicer, W. F. The relation between factor score estimates, image scores, and principal component scores. *Educational and Psychological Measurement*, 1976, *36*, 149–159.

Velicer, W. F. An empirical comparison of the similarity of principal component, image, and factor patterns. *Multivariate Behavioral Research*, in press.

Manuscript received 3/25/74
Revised manuscript received 7/25/75
Final version received 1/16/76