Educational and Psychological Measurement

http://epm.sagepub.com/

Little Jiffy, Mark Iv

Henry F. Kaiser and John Rice

Educational and Psychological Measurement 1974 34: 111
DOI: 10.1177/001316447403400115

EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT
1974, 34, 111-117.

LITTLE JIFFY, MARK IV[1]

HENRY F. KAISER
University of California, Berkeley
and
United States Coast Guard Academy

JOHN RICE
University of California, San Diego

IN this paper three changes and one new development for the method of exploratory factor analysis (a second generation Little Jiffy) developed by Kaiser (1970) are described. Following this short description a step-by-step computer algorithm of the revised method—dubbed Little Jiffy, Mark IV—is presented.

Extensive empirical experience with "a second generation Little Jiffy" (Kaiser, 1970) indicates that the method, for large matrices, consistently mildly underfactors. A revision is called for. Thus, the writers adopt as the answer for the crucially important question of the "number of factors" Guttman's (1954) classic weaker lower bound, the index of the covariance matrix (with zeros in the diagonal) under consideration. This answer is the same as that given by Kaiser's (1956, 1960, 1970) extensively used "eigenvalues greater than one of R."

In a second generation Little Jiffy, for the transformation problem, the writers believe that Kaiser's (1970) winsorizing procedure is unnecessary, and that it would be better to allow possibly a little distortion in the explication of Thurstone's simple pattern and to retain, undisturbed, the elegance (factor-analytic model-free, transformation method-free, column orthogonality) of the Harris-Kaiser (1964) independent cluster solution. The column orthogonality of this solution renders legitimate the interpretation of factors as linear

[1] This research was supported in part by the Office of Computing Activities, National Science Foundation.

combinations of the original variables (or their common parts, or their images) given by the columns of the pattern matrix, in contrast to the typical solution. This column orthogonality also allows one to determine additive relative contributions of factors, in contrast to the typical solution involving correlated factors.

To assess the simplicity of the transformed pattern matrix (and its possible distortion from an ideal simple pattern), an "Index of Factorial Simplicity"

$$IFS = \left[\frac{\sum_j [q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2]}{\sum_j [(q-1)(\sum_s v_{js}^2)^2]} \right]^{1/2} \tag{1}$$

has been developed (Kaiser, 1974), where q is the number of factors and v_{js} is an element in the quartimax transformed Harris (1962) (unit-length) eigenvectors. An analogous $IFS(J)$ for each row follows by removing the Σ_j in the numerator and the denominator of the above formula. Any IFS must lie between zero and one, attaining its maximum value of one only under perfect unifactoriality. Subjective appraisal, based on extensive experience, suggests that these Indices of Factorial Simplicity may be evaluated according to the following table:

in the .90s	marvelous,	
in the .80s	meritorious,	
in the .70s	middling,	(2)
in the .60s	mediocre,	
in the .50s	miserable,	
below .50	unacceptable.	

The above two changes and one new development constitute the revision to a second generation Little Jiffy dubbed "Little Jiffy, Mark III" in an unpublished paper. The following change establishes "Little Jiffy, Mark IV."

Measures of Sampling Adequacy, MSA, as defined by Kaiser (1970), have proved unstable for poor data. Professor Ingram Olkin (personal communication) has suggested that this instability may be corrected by taking a normalized MSA (so that it must lie between zero and one):

$$MSA = \frac{\sum\sum_{j \neq k} r_{jk}^2}{\sum\sum_{j \neq k} r_{jk}^2 + \sum\sum_{j \neq k} q_{jk}^2} \tag{3}$$

overall, and

$$MSA(J) = \frac{\sum_{\substack{k \\ k \neq j}} r_{jk}^2}{\sum_{\substack{k \\ k \neq j}} r_{jk}^2 + \sum_{\substack{k \\ k \neq j}} q_{jk}^2} \tag{4}$$

for each variable separately, where r_{jk} is an original correlation and q_{jk} is an anti-image correlation. Extensive numerical experience with these measures suggests that they also may be evaluated according to the above table (2). The borderline of acceptability here, again .50, occurs when

$$\sum_{j \neq k}\sum q_{jk}^2 = \sum_{j \neq k}\sum r_{jk}^2,$$

where a careful reading of Guttman (1953) indicates that it should be.

This paper appended to Kaiser's (1970) original paper constitutes the final version of Little Jiffy: Little Jiffy, Mark IV.

Computer Algorithm

1. Input and output the correlation matrix R of order p, or input the score matrix Z of order $n \times p$, where n is the number of individuals. R must be positive definite so that n must be greater than p (and also no variable can be a linear combination of other variables). If scores are input, compute and output the vector of means AVE, the diagonal matrix of standard deviations SD, and the correlation matrix R of the variables, and write the original score matrix on auxiliary storage for use at the end of the program in computing factor scores. (If there are missing observations in the score matrix, one wants to insert something in the missing slots to insure that R is positive definite. It is recommended that column means from the available scores be inserted; since most fundamentally below one is working with the image score matrix, one is thus obtaining regression estimates of the missing observations.)

2. Find

$$SS(J) = \sum_{\substack{k \\ k \neq j}} r_{jk}^2; \qquad SS = \sum_{j} SS(J), \tag{5}$$

$$RMS(J) = \sqrt{\frac{SS(J)}{p - 1}}; \qquad RMS = \sqrt{\frac{SS}{p(p - 1)}}, \tag{6}$$

for $j = 1, 2, \cdots, p$, and output $RMS(J)$, the root-mean-square correlation of each variable with the remaining $p - 1$ variables and RMS, the overall root-mean-square correlation.

3. Saving R, find R^{-1}. If R is singular, go to the end of the program; the problem cannot be completed. Let $S^2 = (\text{diag } R^{-1})^{-1}$, the diagonal matrix of anti-image variances. Compute and output $SMC(J) = 1 - s_j^2, j = 1, 2, \cdots, p$, the squared multiple correlation of each variable on the remaining $p - 1$, "best possible" systematic estimates of communalities (Guttman, 1956). Write R^{-1} on auxiliary storage for use in computing the factor score weight matrix at the end of the program.

4. Find the anti-image correlation matrix $Q = SR^{-1}S$ and then determine

$$SSS(J) = \sum_{\substack{k \\ k \neq j}} q_{jk}^2; \qquad SSS = \sum_{j} SSS(J). \tag{7}$$

5. Now compute

$$MSA(J) = \frac{SS(J)}{SS(J) + SSS(J)}; \qquad MSA = \frac{SS}{SS + SSS} \tag{8}$$

for $j = 1, 2, \cdots, p$. $MSA(J)$ is the revised (Mark IV) Measure of Sampling Adequacy for the jth variable, and MSA is the revised (Mark IV) overall Measure of Sampling Adequacy.

6. Find the eigenvalues of R and let q equal the number of eigenvalues greater than one; q is the number of factors.

7. Form Harris' (1962) covariance matrix, $C = S^{-1}RS^{-1}$, and then find the diagonal matrix M^2 of the q largest eigenvalues (Harris eigenvalues) of C and the associated q unit-length column (Harris) eigenvectors. Let E be the $p \times q$ matrix of these eigenvectors.

8. Saving a copy of E, apply the (raw) quartimax criterion (Wrigley and Neuhaus, 1954) to E, calling the result V. An algorithm for the quartimax method is given by three changes in Kaiser's (1959) algorithm for the varimax criterion. In that write-up (a) omit steps 2 and 17; (b) in step 5 let $\tan 4\phi = A/B$; and (c) tighten the convergence criterion in step 11 to at least $\delta \leq 10^{-4}$.

9. For each column of V find the sum of cubes and, if the sum is negative, change the signs of all the entries in the column.

10. Find $T = E'V$, the quartimax transformation matrix.

11. Replace the diagonal elements of the diagonal matrix M^2 of Harris eigenvalues with

$$\frac{(m_s^2 - 1)^2}{m_s^2} \tag{9}$$

for $s = 1, 2, \cdots, q$. Call this new diagonal matrix M^2, also. It contains the Harris eigenvalues of the Harris rescaled image covariance matrix.

12. Form $LSTAR = T'M^2T$. $LSTAR$ is the factor intercovariance matrix. Let $D^2 = \text{diag}\ (LSTAR)$ be the diagonal matrix of "natural" factor variances.

13. Reorder elements of D^2 in order of decreasing size, and reorder columns of V and rows and columns of $LSTAR$ accordingly.

14. From V find the $IFS(J)$ and IFS

$$IFS(J) = \left[\frac{q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2}{(q - 1)(\sum_s v_{js}^2)^2}\right]^{1/2} \tag{10}$$

$$IFS = \left[\frac{\sum_j [q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2]}{\sum_j [(q - 1)(\sum_s v_{js}^2)^2]}\right]^{1/2} \tag{11}$$

for $j = 1, 2, \cdots, p$.

15. Multiply each element of V by $\sqrt{p}$—a step thus changing the loadings from column-normalized to column-standardized. Output this column-standardized factor pattern matrix. Return it to its original column-normalized version V by dividing each element by $\sqrt{p}$. Now output $IFS(J)$ and IFS.

16. For each factor compute

$$\frac{100\ d_s^2}{\sum_s d_s^2} \tag{12}$$

for $s = 1, 2, \cdots, q$, and output these relative variance contributions of factors (percentages).

17. Find the factor intercorrelation matrix,

$$L = D^{-1}(LSTAR)D^{-1}, \tag{13}$$

and output

18. a. Form $A = SVD$, the factor pattern matrix (conventionally scaled).

b. Form and output $B = AL$, the factor structure matrix.

c. Output A, asterisking the salients—those elements of A which are larger than one in the column-standardized factor pattern matrix V.

19. Calculate the standardized factor score weight matrix W:

$$W' = D^{-2}A'(S^{-2} - R^{-1}) \tag{14}$$

and output. (Note that $(S^{-2} - R^{-1})$ is just $-R^{-1}$ with zeros in the diagonal.)

20. Determine the domain validities (indices of reliability) of the factors, DV(S):

$$DV(S) = \left[\left(\frac{p}{p-1}\right)\left(1 - \sum_i w_{is}^2\right)\right]^{1/2} \tag{15}$$

for $s = 1, 2, \cdots, q$, and output.

If a correlation matrix was input at the beginning, the problem is finished. If scores were input, the program goes on to compute (image) factor scores. The formulas used below will be developed in detail in a forthcoming paper.

21. Rescale the standardized factor score weight matrix W to be

$$(100)W'(SD)^{-1}$$

where (SD) is the diagonal matrix of original standard deviations. Call this new factor score weight matrix W, also. It has been rescaled to accept the original raw scores and to produce factor scores with standard deviations equal to 100.

22. To calculate factor scores, one goes through the following loop for each individual:

a. Read in a row vector z' of p (raw) scores for a given individual.
b. Subtract the vector of original means AVE from z. Call these centered scores z, also.
c. Let $x = W'z$.
d. Find $x + 500$, the desired factor scores for one individual. These factor scores have mean 500 for all factors.
e. Print and punch the vector of factor scores for one individual.
f. Go back to (a) for the next individual's observation vector.

REFERENCES

Guttman, L. Image theory for the structure of quantitative variates. *Psychometrika,* 1953, 18, 277–296.

Guttman, L. Some necessary conditions for common factor analysis. *Psychometrika,* 1954, 19, 149–161.

Guttman, L. Best possible systematic estimates of communalities. *Psychometrika*, 1956, 21, 273–285.

Harris, C. W. Some Rao-Guttman relationships. *Psychometrika*, 1962, 27, 247–263.

Harris, C. W. and Kaiser, H. F. Oblique factor analytic solutions by orthogonal transformations. *Psychometrika*, 1964, 29, 347–362.

Kaiser, H. F. The varimax method of factor analysis. Unpublished doctoral dissertation, University of California, Berkeley, 1956.

Kaiser, H. F. Computer program for varimax rotation in factor analysis. EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 1959, 19, 413–420.

Kaiser, H. F. The application of electronic computers to factor analysis. EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 1960, 20, 141–151.

Kaiser, H. F. A second generation Little Jiffy. *Psychometrika*, 1970, 35, 401–415.

Kaiser, H. F. An index of factorial simplicity. *Psychometrika*, in press, 1974.

Wrigley, C. W. and Neuhaus, J. O. The quartimax method: An analytical approach to orthogonal simple structure. *British Journal of Statistical Psychology*, 1954, 7, 81–91.