DETERMINING SIGNIFICANT PRINCIPAL COMPONENTS: A PROBABILITY TEST FOR EIGENVALUES

Raymond Hubbard, SUNY Fredonia
Vinay Pandit, Saint Bonaventure University

Abstract

When employing factor analysis a major problem confronting researchers is the determination of how many components to retain for subsequent investigation. Many marketing scholars address this issue by automatically invoking the Kaiser-Guttman, or eigenvalue-greater-than-one, procedure. Yet this procedure is not without certain limitations. The authors therefore propose an alternative measure, a probability test for the significance of eigenvalues, based upon a random intercepts model. Preliminary results of the application of the model on marketing and psychology data are presented.

Introduction

Despite the variety of multivariate techniques available to the marketing researcher, factor analytic procedures remain among the most commonly utilized. The popularity of factor analysis as a marketing tool has, however, been tainted by a number of problematical issues. Certain of these issues concern the inappropriate application of factor analysis in assorted marketing studies (Stewart 1981), while others revolve around the development and refinement of factor analytic methods per se (Acito and Anderson 1980; Acito, Anderson and Engledow 1980). With regard to the latter, Stewart (1981, p. 58) has remarked that: "Perhaps no problem has generated more controversy and misunderstanding than the number of factors problem." This view is shared by others (e.g., Hakstian and Muller 1973). The essence of this problem may be summarized by asking the following related questions: How many factors should be extracted from a correlation matrix? When should factoring cease? What is the "best" or "correct" number of factors to retain for further analysis and rotation? Such questions acquire great importance because of their obvious theoretical implications, and because of the fact that rotation is not particularly robust to the number of initial factors obtained.

Principal components and factor analysis are frequently thought of as being similar insofar as both methods are concerned with data reduction. They are different however. Factor analysis presupposes the existence of an underlying hypothetical causal model, and as such is theoretically capable of prespecifying the number of factors to be extracted. Principal component analysis, on the other hand, is more empirically oriented and focuses upon the linear transformation of a set of observed data. This paper is concerned with principal component analysis.

In the absence of powerful marketing (or other) theories which would define the number of factors to be extracted, a series of tests and rules of thumb have understandably been formulated to address this critical issue. By far the most routinely employed criterion for factor extraction is the Kaiser-Guttman (or eigenvalue-greater-than-one rule, roots criterion) rule. This rule, together with a number of its drawbacks, is discussed below.

The Kaiser-Guttman Rule

The Kaiser-Guttman (hereafter KG) rule is so ubiquitous that it is usually incorporated as an automatic default option in most statistical packages. This rule is identified with Kaiser's (1960; 1970; 1974) so-called "Little Jiffy." Little Jiffy describes a factor analytic approach which proceeds with a principal component analysis of a correlation matrix, followed by the retention of components whose eigenvalues exceed or are equal to unity, and whose final solution is determined by varimax rotation. Based on a survey of JMR articles published between November 1964 and February 1978, in which factor analysis was the predominant investigative tool, Acito and Anderson (1980) were able to conclude that Little Jiffy was the most frequently adopted factor analytic strategy.

The logic underlying the KG criterion can be explained by considering a correlation matrix of full rank N which is being analyzed. In general, this matrix would yield N principal components and the sum of the eigenvalues associated with these components would be equal to the trace of the correlation matrix. If λ_i is the eigenvalue of the i^{th} principal component, it can be demonstrated that λ_i/N is the proportion of the variance in the standardized data that can be explained by the principal component. Even if the principal components are produced by random grouping of the variables, each component on an average would be expected to account for 1/N of the variance in the data. It would therefore have an eigenvalue of one. And so, if the eigenvalue of a factor is equal to unity, it contributes its "average" share to the explanation of the variance in the data. Such factors, together with those whose eigenvalues exceed unity, are therefore considered to be of explanatory significance.

Notwithstanding the popularity of the KG rule, it should be noted that several limitations accompany its usage. First, the choice of one or greater as the cut-off point for component retention is too arbitrary a criterion. In any principal component analysis a score of one is by definition the average value of all the eigenvalues involved. It therefore follows that certain eigenvalues will necessarily exceed unity, while others will not. Consequently, some eigenvalues greater than unity must occur irrespective of the "significance," in either a statistical or a substantive sense, of their associated components. Rather than employing a criterion that forces acceptance of some components as significant and others as not, it appears preferable to develop a measure which is

not constrained in this fashion.

Second, the eigenvalue rule is based on the assumption that a population correlation matrix is being factored (Stewart 1981). Otherwise expressed, no sampling errors exist in the data being analyzed. But the majority of marketing studies are conducted on the basis of samples, and are thereby subject to sampling errors in the correlation matrix which will be expected to distort the resultant eigenvalues. Therefore the significance of eigenvalues should be assessed via probabilistic tests, which the KG criterion does not provide.

Third, prompted by continuing advances in factor analytic methodology, Kaiser himself has acknowledged the need to modify Little Jiffy. He comments that (Kaiser 1970, p. 403): "Little Jiffy, that seemingly reliable workhorse for exploratory studies, is old, and in need of an overhaul." Fourth, Stewart (1981), citing various studies from the psychology literature, maintains that the most damaging criticism leveled against the KG rule originates from Monte Carlo work with the method. More specifically, he notes that this criterion is at best an approximate one which frequently generates an exorbitant number of factors.

It is clear that room for optional procedures to assist with the number of factors problem currently exist. We therefore propose an alternative test to the KG criterion. The test itself, which we feel has many beneficial qualities, is derived from the random intercepts or random partitions model (Feller 1966; Pandit 1978).

Random Intercepts Model

In some cases the random intercepts (henceforth RI) test may supplant the roots criterion, while in others it may supplement it. Regardless, the test is sufficiently versatile such that it can readily be employed in conjunction with the KG rule should this be desired. Viewed in this manner, application of the RI test is perhaps consonant with requests that Little Jiffy requires some modifications (Acito and Anderson 1980; Kaiser 1970). Again, the RI approach, concerned as it is with principal component analysis, appears well-suited to mainstream marketing factor analytic situations (Green 1978). Finally, the RI model is flexible enough to accommodate aspects of the under-over factoring controversy by allowing the researcher reasonable latitude to specify different significance levels for component extraction.

The relationship between principal components analysis and the RI model becomes clear when we consider the following null hypothesis:

H_o: Principal components analysis generates a random grouping of variables.

That is, eigenvalues occur as a result of the functioning of random phenomena. Note: although the average eigenvalue will still be one, as in the roots criterion, we pay explicit attention to the distribution of the eigenvalues. This distribution follows from the operation of the RI model.

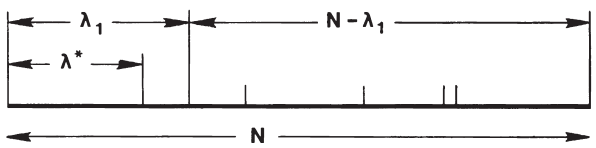
H_1: Principal components analysis yields a grouping of variables that is not the result of random processes.

In other words, the magnitude of an eigenvalue should be sufficiently large so as to preclude random grouping. The discussion below is concerned with the development of the random intercepts model itself.

Consider a line of length N portraying the sum of the eigenvalues connected with a matrix of rank N (**Figure 1**). N-1 random and independent intercepts arrayed on this line will result in segments of length λ_i which would represent eigenvalues associated with random components ($N = \sum_{i=1}^{N} \lambda_i$). Recall that the eigenvalue of a component is equal to the variance accounted for by that component. Therefore, our examination of eigenvalues is formally equivalent to an investigation of the variance in the data.

As **Figure 1** illustrates, the leftmost segment (or eigenvalue) is assumed to be of magnitude λ_1, and will be considered to be significant if it is longer than some predetermined value λ^*. λ^* itself, which is derived below, is defined as the smallest significant eigenvalue.

FIGURE 1
HYPOTHETICAL RANDOM PARTITIONING
OF THE SUM OF EIGENVALUES



If the intercepts are generated at random, with uniform distribution over N, the probability that a cut appears to the left of λ_1 is denoted by λ_1/N. Similarly, the probability that any intercept would fall to the right of λ_1 would then be $(N-\lambda_1)/N$. Since λ_1 is the length of the leftmost segment, all N-1 intercepts will necessarily be to the right of λ_1. Furthermore, it follows that the probability that all intercepts will lie to the right of λ_1 is given by:

$$P(\lambda_1,N) = \left(\frac{N-\lambda_1}{N}\right)^{N-1} \quad (1)$$

And (1) is also synonymous with the probability that the leftmost segment (eigenvalue) is at least of length λ_1.

In order to accept this length, λ_1, as significant it should be greater than or equal to some prespecified value λ^*, where:

$$P(\lambda^*,N) = \left(\frac{N-\lambda^*}{N}\right)^{N-1} = \alpha \quad (2)$$

Alpha (α) would be a probability specification similar to a level of significance employed in conventional statistical tests. Given values of α (the level of significance) and N (the number of segments/eigenvalues generated), we can compute the test length λ^* for comparison with the generated length λ_1. If $\lambda_1 \geq \lambda^*$, we accept the length of the segment/eigenvalue as significant. We therefore solve (2) in order to obtain our values for λ^*:

$$\lambda^* = N\,(1-\alpha^{\frac{1}{N-1}}) \qquad (3)$$

Using (3) we calculate values of λ^* for various levels of α and N. These values are listed in Table 1.

The distribution of segment proportions (λ/N) resulting from N-1 random intercepts on a line of length N can be shown to be of the form:

$$f(\lambda/N,N) = (N-1)(1-\frac{\lambda}{N})^{N-2} \qquad (4)$$

Preliminary Testing of the RI Rule

In order to illustrate how component extraction via the RI rule compares with the KG criterion, the authors conducted preliminary tests by subjecting eight data sets to a complete principal components analysis. Four of these sets were from marketing sources (Hornik 1980; Mukherjee 1971; Perreault, Jr. and Russ 1976; Stephenson 1963) and four were

TABLE 1
SIGNIFICANT EIGENVALUES (λ^*)[a]

Number of Variables	Probabilities											
	0.500	0.400	0.300	0.200	0.100	0.050	0.040	0.030	0.020	0.010	0.005	0.001
2	1.001	1.201	1.401	1.601	1.801	1.901	1.921	1.941	1.961	1.981	1.991	1.999
4	0.826	1.053	1.323	1.661	2.144	2.527	2.633	2.758	2.915	3.139	3.317	3.601
6	0.777	1.005	1.284	1.652	2.215	2.705	2.849	3.025	3.257	3.612	3.921	4.493
8	0.755	0.982	1.265	1.644	2.243	2.786	2.949	3.153	3.426	3.857	4.248	5.018
10	0.742	0.968	1.253	1.638	2.258	2.832	3.007	3.227	3.526	4.006	4.450	5.359
20	0.717	0.942	1.229	1.625	2.283	2.918	3.117	3.371	3.722	4.305	4.868	6.097
30	0.709	0.934	1.220	1.620	2.290	2.945	3.152	3.417	3.786	4.405	5.010	6.359
40	0.705	0.929	1.216	1.618	2.294	2.958	3.169	3.440	3.818	4.456	5.082	6.493
50	0.703	0.927	1.214	1.616	2.296	2.966	3.180	3.454	3.837	4.486	5.125	6.575
60	0.701	0.925	1.212	1.615	2.297	2.971	3.186	3.463	3.850	4.506	5.154	6.630
70	0.700	0.924	1.211	1.614	2.298	2.975	3.191	3.469	3.859	4.520	5.174	6.669
80	0.699	0.923	1.210	1.614	2.299	2.977	3.195	3.474	3.866	4.531	5.190	6.699
90	0.699	0.922	1.210	1.613	2.299	2.979	3.197	3.478	3.871	4.539	5.202	6.722
100	0.698	0.922	1.209	1.613	2.300	2.981	3.200	3.480	3.875	4.546	5.212	6.740
200	0.696	0.919	1.207	1.612	2.301	2.989	3.210	3.494	3.894	4.576	5.255	6.824
300	0.695	0.918	1.206	1.611	2.302	2.991	3.213	3.498	3.900	4.586	5.270	6.852
400	0.695	0.918	1.206	1.611	2.302	2.992	3.214	3.500	3.903	4.591	5.277	6.866
500	0.695	0.918	1.205	1.611	2.302	2.993	3.215	3.502	3.905	4.594	5.281	6.874
1000	0.694	0.917	1.205	1.610	2.303	2.995	3.217	3.504	3.909	4.600	5.290	6.891

[a]Accept an eigenvalue as significant if it is equal to, or exceeds, λ^*.

taken from the psychology literature (Davis 1944; Fleishman and Hempel Jr. 1954; Harman 1976; Karlin 1941). At this juncture no claim is made that these initial results are necessarily generalizable. Further testing is obviously warranted before such conclusions can be offered. Table 2 presents the findings of the eight principal components analyses.

For purely illustrative purposes consider the second data set (Harman 1976). Consulting Table 2 we find in this case that three components, collectively accounting for some 98 percent of the variance in the data, are identified when the KG rule is invoked. To ascertain how many components are considered to be significant when adopting the RI rule, we refer to Table 1. With twenty variables and a desired significance level of, say, .01, an eigenvalue of magnitude 4.305 must be obtained. The first component, with an associated eigenvalue of 12.62, is clearly significant at this level. Similarly the second and third components are found to be significant at the .02 and .05 levels respectively. Both the KG and RI rules produced three significant components in this particular instance.

However, further inspection of Table 2 reveals that, in general, there are differences between the KG and RI rules with respect to the number of components each method generates. Compared with the KG criterion, the RI rule displays a tendency to underfactor correlation matrices. Whereas on average the KG rule elicits 2.5 components per study, the RI method produces only 1.5. Furthermore, in empirical research the weight of opinion tends to err in favor of the overfactoring of a matrix by one or two components should the under- over factoring controversy emerge (Cattell 1952; Stewart 1981).

Nevertheless, a number of points are in order. First, there is nothing sacred about the KG rule, and we have cited earlier a number of shortcomings connected with it. Second, the predominant usage of the KG rule in empirical works in marketing and psychology frequently appears to be of a nescient character. As Cattell and Vogelmann (1977, pp. 322-23) note:

"The practice of building the KG into factoring programs, because of its extreme cheapness, and proceeding with no break for the investigator to apply other checks for number of factors (leaving the novice even unaware of the basis of the decision) is quite indefensible."

Third, to say that the RI rule appears to underfactor is only to say that it underfactors in comparison with the KG criterion. Otherwise expressed "underfactoring" is a relative term. Indeed there is evidence to indicate that the KG rule is susceptible to charges of overfactoring (Cattell and Vogelmann 1977; Hakstian, Rogers, and Cattell 1982; Stewart 1981). This suggests that a "conservative" test such as that embodied in the RI model is not without its merits. Moreover, since the logic underlying the KG and RI rules is different, we should anticipate that the two approaches will yield different results.

TABLE 2
PRINCIPAL COMPONENTS ANALYSES OF EIGHT DATA SETS: KG vs. RI

Data Set	N	n	# of Components Extracted KG	(Eigenvalues ≥ 1)	RI	RI/KG	% Variance "Explained" KG	RI
Fleishman/Hempel, Jr.	26	197	4	10.99***, 2.97* 1.40, 1.09	2	0.5	63.3	42.3
Harman	20	60	3	12.62***, 4.04††, 2.98*	3	1.0	98.2	98.2
Karlin	19	120	5	4.37**, 2.93* 1.37, 1.14, 1.06	2	0.4	56.9	38.2
Mukherjee	14	94	1	10.64***	1	1.0	76.0	76.0
Hornik	11	145	2	7.39***, 1.50	1	0.5	88.9	73.9
Perreault/Russ	10	216	2	5.19**, 1.16	1	0.5	63.5	51.9
Stephenson	10	38	2	4.04**, 1.50	1	0.5	55.4	40.4
Davis	9	421	1	5.20***	1	1.0	57.7	57.7

N = variable numbers; n = sample size. ***, **, ††, and *, indicates significance at the .001, .01, .02, and .05 levels, respectively.

Fourth, the RI approach seems to yield parsimonious results. Thus when employing the KG rule it takes, on an average over all eight studies, 2.5 components to account for seventy percent of the variance in the data. On the other hand the RI rule, which extracts an average of only 1.5 components per study, nevertheless explains some sixty percent of data variation. That is, the RI approach picks up dominant eigenvalues and their components rather than peripheral ones. For example, in the Fleishman and Hempel, and Karlin studies, eigenvalues of 1.09 and 1.06 respectively are considered significant according to the KG procedure. These eigenvalues are ignored by the RI rule. Thus it is worthwhile reiterating that the KG criterion will accept any eigenvalue as significant if its magnitude exceeds unity, that is the average value of all the eigenvalues involved in the study. In essence we have developed a procedure which insures that it is not sufficient to exceed a given average value, but rather that one should significantly exceed this value. And we have anchored this procedure in probabilistic principles. The advantage of adopting the RI rule is that it provides a probability level for significant component retention that the KG criterion cannot match.

Conclusions

We feel that the logic embodied in the RI model, originally developed by Pandit (1978), represents an appropriate alternative to the KG criterion. By considering the probability of obtaining certain magnitudes of eigenvalues, the RI model supplies a test for determining the significance of eigenvalues, and thus a more rational justification for identifying significant principal components. A word of caution is appropriate at this juncture. The RI test suffers from the fact that it is not based on any error theory or sampling distribution. This problem is currently being addressed. In any event it appears reasonable to suggest that, whether used separately or in concert with the KG rule, the RI alternative warrants further attention in marketing studies.

References

Acito, Franklin and Ronald D. Anderson (1980), "A Monte Carlo Comparison of Factor Analytic Methods," Journal of Marketing Research, 17 (May), 228-36.

_____, _____, and Jack L. Engledow (1980), "A Simulation Study of Methods for Hypothesis Testing in Factor Analysis," Journal of Consumer Research, 7 (September), 141-50.

Cattell, Raymond B. (1952), Factor Analysis. New York: Harper and Brothers.

_____, and S. Vogelmann (1977), "A Comprehensive Trial of the Scree and KG Criteria for Determining the Number of Factors," Multivariate Behavioral Research, 12 (July), 289-325.

Davis, Frederick B. (1944), "Fundamental Factors of Comprehension in Reading," Psychometrika, 9 (September), 185-97.

Feller, William (1966), An Introduction to Probability Theory and Its Applications. New York: John Wiley and Sons, Inc.

Fleishman, Edwin A. and Walter E. Hempel, Jr. (1954), "Changes in Factor Structure of a Complex Psychomotor Test as a Function of Practice," Psychometrika, 19 (September), 239-52.

Green, Paul E. (1978), Analyzing Multivariate Data. Hinsdale, Illinois: Dryden Press.

Hakstian, A. Ralph and Victor J. Muller (1973), "Some Notes on the Number of Factors Problem," Multivariate Behavioral Research, 8 (October), 461-75.

_____, W. Todd Rogers, and Raymond B. Cattell (1982), "The Behavior of Number-of-Factors Rules with Simulated Data," Multivariate Behavioral Research, 17 (April), 193-219.

Harman, Harry H. (1976), Modern Factor Analysis. Chicago: University of Chicago Press.

Hornik, Jacob (1980), "Quantitative Analysis of Visual Perception of Printed Advertisements," Journal of Advertising Research, 20 (December), 41-8.

Kaiser, Henry F. (1960), "The Application of Electronic Computers to Factor Analysis," Educational and Psychological Measurement, 20 (Spring), 141-51.

_____ (1970), "A Second Generation Little Jiffy," Psychometrika, 35 (December), 401-15.

_____, and John Rice (1974), "Little Jiffy Mark IV," Educational and Psychological Measurement, 34 (Spring), 111-17.

Karlin, J. E. (1941), "Music Ability," Psychometrika, 6 (February), 61-5.

Mukherjee, Bishwa N. (1971), "A Factor Analysis of Some Qualitative Attributes of Coffee," in Multivariate Analysis in Marketing: Theory and Application. David A. Aaker (ed.), Belmont, California: Wadsworth Publishing Co., 245-51.

Pandit, Vinay Y. (1978), "Data Breeding: Estimation of Micro-Data in an Incomplete Data-Base." Unpublished Ph.D. Dissertation, Graduate School of Business Administration, Columbia University.

Perreault, Jr., William D. and Frederick A. Russ (1976), "Physical Distribution Service in Industrial Purchase Decisions," Journal of Marketing, 40 (April), 3-10.

Stephenson, William (1963), "Public Images of Public Utilities," Journal of Advertising Research, 3 (December), 34-9.

Stewart, David W. (1981), "The Application and Misapplication of Factor Analysis in Marketing Research," Journal of Marketing Research, 18 (February), 51-62.