PRINCIPAL COMPONENTS, RANDOM NUMBERS, AND TRANSFORMATIONS

Charles D. Dziuban and Edwin C. Shirkey/Florida Technological University

Introduction

Shaycoft (1970) illustrated an important problem that would be encountered if one applied the method of principal components to a specially constructed correlation matrix. That data set consisted of 10 variables of interest (from Project Talent) combined with 4 random deviates. The 14 variable mixed set has been more completely described elsewhere (Dziuban and Harris, 1973). Shaycoft's (1970) varimax transformed pattern would have forced one to make an interpretation of random variables as the basis of a meaningful component. In their discussion of that matrix, Dziuban and Harris (1973) demonstrated that the common factor and image procedures would correctly alert the investigator to the random variables. Later Dziuban and Shirkey (1974A, 1974B) and Dziuban, Shirkey, and Peeples (1975) explored other procedures for identifying those suspect variates.

The question may arise, however, with the principal components solution as to whether transformational techniques other than normal varimax would result in an acceptable solution, i.e., the random deviates exhibiting essentially zero pattern coefficients. Since there are those (Velicer, 1977) who feel that factor analysis, image analysis, and principal components will yield similar results when applied to common data sets, it was sought in this study to determine whether or not an acceptable solution might be derived through the use of principal components. Specifically, several blind rotational schemes were applied to the Shaycoft (1970) data in an attempt to eliminate the random numbers from the pattern matrix. It has been common practice (Hakstian, 1971 and Hakstian and Abell, 1974) to apply several rotational schemes to the same data. To recap, we know that the principal components solution is wrong. An attempt was made through orthogonal and oblique transformation to make it correct.

Procedures

The Shaycoft (1970) matrix was subjected to a principal components resolution. Raw components were extracted according to the eigenvalues greater than one ($\lambda > 1$). The raw components were transformed according to the following criteria.

Orthogonal Procedures

Quartimax. This method is often credited to Neuhaus and Wrigley (1954), but Carroll (1953), Saunders (1953), and Ferguson (1954) working independently arrived at identical techniques. The practical objective of the procedure is to minimize the complexity of each variable; that is, a variable should exhibit a substantial pattern coefficient on one factor and near zero values on all others.

Equimax. This is another of the often cited but rarely used methods of orthogonal transformation. Saunders (1962) sought to combine the varimax and quartimax procedures in an attempt to salvage the best features of each. He discovered the varimax method involved the essential terms of quartimax plus some additional ones. Essential to both computational procedures was a value K. Saunders found that a wide range of acceptable solutions could be derived with changing values of K once it exceeded some minimum value.

Oblique Procedures

Direct Oblimin. The direct oblimin procedure was designed in part by Jennrich and Sampson (1966) to avoid working with reference axes in favor of the primary pattern coefficients. The procedure seeks a simple structure solution by minimizing a function of the primary pattern coefficients.

Promax. Hendrickson and White (1964) devised a method of procrustean application in the blind transformation circumstance. In the promax method, a preliminary orthogonal solution, usually varimax, is transformed to an oblique solution through a least squares fit to a target matrix in which the elements are those of the original matrix raised to some power greater than one with the signs of the original coefficients being retained. Hendrickson and White (1964) began with the assumption that the original orthogonal solution was usually a fairly good approximation to the finally obtainable simple structure solution. By using the powered elements of the orthogonal solution as the basis of a target matrix for an unrestricted procrustean transformation, they accentuated the division between the salient and nonsalient coefficients. Hendrickson and White (1964) recommended that the power to which the original coefficients be raised should be either two or four.

Harris-Kaiser Orthoblique

In the general Harris-Kaiser (1964) orthoblique method, the primary pattern matrix P is obtained by:

$$P = WQL^{C}TD$$

Harris and Kaiser (1964) identified two special cases. In the general formulation:

C = 0 The independent cluster solution

C = .5 P′P proportional to ϕ the factor correlation matrix

As C increases to 1, the solution approaches orthogonality. Best results are usually obtained for factorially simple data with C = 0 and with more complex data C = 1 usually yields the best results.

Results

The original Shaycoft correlation matrix may be found in Dziuban and Harris (1973). All the correlations among the 10 tests of interest and random deviates were zero to the first place as they were among the random deviates. The normal varimax solution (Dziuban and Harrix, 1973) is not reported here but failed to produce an acceptable solution.

The results of the quartimax and equimax solutions are presented in Table 1. It may be observed that both transformational techniques

TABLE 1

Quartimax

	1	2	3
1	86	-02	-04
2	79	02	-07
3	87	00	00
4	85	003	-02
5	59	-04	01
6	77	03	05
7	84	03	-04
8	60	01	09
9	71	03	08
10	83	00	01
11	-03	$\underline{-68}$	-06
12	03	-18	$\underline{72}$
13	-02	19	$\underline{69}$
14	00	$\underline{70}$	-05

Equimax

	1	2	3
1	86	-01	-04
2	79	-02	-06
3	87	01	00
4	85	-02	-01
5	00	-04	02
6	77	04	-05
7	84	04	-05
8	61	02	10
9	71	04	08
10	83	01	02
11	-02	$\underline{-68}$	-06
12	03	-18	$\underline{72}$
13	-03	18	$\underline{69}$
14	00	$\underline{70}$	-05

yielded virtually identical results. The first components were dominated by the Project Talent variables while the second components were associated with random variates 11 and 14. The third components were most highly correlated with random variables 12 and 13.

The results of the oblique procedures are presented in Table 2.

TABLE 2

Oblimin Delta = 0

	1	2	3
1	86	-02	04
2	79	01	-07
3	87	00	-00
4	85	-03	-02
5	60	-05	01
6	77	03	-05
7	84	03	-04
8	60	01	09
9	71	03	07
10	83	00	01
11	-02	$\underline{-68}$	-05
12	03	18	$\underline{72}$
13	-03	18	$\underline{68}$
14	00	$\underline{70}$	-06

Correlations

1.00		
.01	1.00	
.01	.01	1.00

Promax-Powering Factor = 2

	1	2	3
1	86	-31	-15
2	79	03	-05
3	86	02	01
4	85	-02	00
5	59	-03	03
6	77	04	-04
7	84	05	-02
8	60	02	11
9	71	04	09
10	83	02	03
11	01	$\underline{-68}$	-03
12	00	-20	$\underline{73}$
13	-05	17	$\underline{67}$
14	-02	$\underline{71}$	-08

Correlations

1.00		
.01	1.00	
.02	.07	1.00

Promax-Powering Factor = 4

	1	2	3
1	89	-29	-14
2	79	05	-04
3	86	03	01
4	85	03	00
5	59	-02	04
6	76	06	-03
7	84	06	-02
8	59	03	10
9	70	05	09
10	82	03	03
11	02	$\underline{-68}$	-01
12	-02	-21	$\underline{73}$
13	10	15	$\underline{67}$
14	-04	$\underline{71}$	-10

Correlations

1.00		
.03	1.00	
.07	.12	1.00

Harris-Kaiser P = 0

	1	2	3
1	84	-04	-17
2	79	00	-06
3	86	-01	-00
4	85	-04	-01
5	59	-05	02
6	77	02	-05
7	84	01	-04
8	61	00	09
9	71	02	08
10	83	-00	01
11	06	-68	-03
12	04	-19	$\underline{73}$
13	00	18	$\underline{68}$
14	04	$\underline{71}$	-08

Correlations

1.00		
-.04	1.00	
-.03	.05	1.00

Harris-Kaiser P = .5

	1	2	3
1	87	32	-17
2	79	02	06
3	86	01	00
4	85	-02	-01
5	60	-04	02
6	76	03	-04
7	84	04	-03
8	60	01	10
9	71	03	08
10	82	01	02
11	00	$\underline{-68}$	-05
12	03	-18	$\underline{72}$
13	03	17	$\underline{68}$
14	03	$\underline{70}$	-06

Correlations

1.00		
.02	1.00	
.00	.03	1.00

It may be readily observed that all of the solutions yielded virtually identical results which were, in turn, highly similar to the orthogonal solutions—the one general and two random components. All component correlation matrices revealed the components to be, in fact, orthogonal.

Discussion

We have provided a demonstration of several "blind" orthogonal and oblique transformations to a specially mixed set of variables. In the initial solution presented by Shaycoft (1970)—principal components with varimax little jiffy—something very understandable occurred. The first component was put through the centroid of the 10 Project Talent variables. The random variables which were independent of the Talent set are orthogonally located with respect to them. Consequently, the second component, which must be uncorrelated with the first, is put through random variables 11 and 14. The third orthogonal component is positioned through variables 12 and 13. These results are to be expected because of the nature of the principal component solution. It is a procedure for deriving uncorrelated variables within the variable space. The procedure is not common factor analysis. Therefore, the random variables, for which no uniqueness has been taken into account, command a salient position on the components. None of the transformational techniques, orthogonal or oblique, placed the components in close proximity since they are, in fact, positioned orthogonally with respect to each other.

References

Carroll, S. B. "An Analytic Solution for Approximating Simple Structure in Factor Analysis." *Psychometrika,* 1953, *18,* 23-38.

Dziuban, C. D., and Harris, C. W. "On the Extraction of Components and the Applicability of the Factor Model." *American Educational Research Journal,* 1973, *10,* 93-99.

Dziuban, C. D., and Shirkey, E. C. "On the Psychometric Assessment of Correlation Matrices." *American Educational Research Journal,* 1974A, *11,* 211-216.

Dziuban, C. D., and Shirkey, E. C. "When Is a Correlation Matrix Appropriate for Factor Analysis: Some Decision Rules." *Psychological Bulletin,* 1974B, *81,* 358-361.

Dziuban, C. D.; Shirkey, E. C.; and Peeples, T. O. "On the Independence of Variable Sets." *Educational and Psychological Measurement,* 1975, *35,* 539-540.

Ferguson, G. A. "The Concept of Parsimony in Factor Analysis." *Psychometrika,* 1954, *19,* 281-290.

Hakstian, A. R. "A Comparative Evaluation of Several Prominent Methods of Oblique Factor Transformation." *Psychometrika,* 1971, *36,* 175-193.

Hakstian, A. R., and Abell, R. A. "A Further Comparison of Oblique Factor Transformation Methods." *Psychometrika,* 1974, *39,* 429-444.

Harris, C. W., and Kaiser, H. F. "Oblique Factor Analytic Solutions by Orthogonal Transformations." *Psychometrika,* 1964, *29,* 347-362.

Hendrickson, A. E., and White, P. O. "PROMAX: A Quick Method for Rotation to Oblique Simple Structure." *British Journal of Statistical Psychology,* 1964, *17,* 65-70.

Jennrich, R. I., and Sampson, P. F. "Rotation for Simple Loadings." *Psychometrika,* 1966, *31,* 313-323.

Neuhaus, J. O., and Wrigley, C. "The Quartimax Method: An Analytical Approach to Orthogonal Simple Structure." *British Journal of Statistical Psychology,* 1954, *7,* 81-91.

Saunders, D. R. "An Analytic Method for Rotation to Orthogonal Simple Structure." *Research Bulletin 53-10.* Princeton, New Jersey: Educational Testing Service, 1953.

Saunders, D. R. "Trans-Varimax: Some Properties of the Ratiomax and Equimax for Blind Orthogonal Rotation." Paper presented at the Annual Meeting of the American Psychological Association, St. Louis, 1962.

Shaycoft, M. F. "The Eigenvalue Myth and the Data Reduction Fallacy." Paper presented at the Annual Meeting of the American Education Research Association, Minneapolis, March, 1970.

Shirkey, E. C., and Dziuban, C. D. "A Note on Some Sampling Characteristics of the Measure of Sampling Adequacy (MSA)." *Multivariate Behavioral Research,* 1976, *11,* 125-128.

Velicer, F. "An Empirical Comparison of the Similarity of Principal Component Image and Factor Patterns." *Multivariate Behavioral Research,* 1977, *12,* 3-22.