₹ 200 ISSN - 2249-555X

Volume : 1 Issue : 6 March 2012

Indian Journal Of Applied Research

Journal for All Subjects

www.ijar.in

Listed in International ISSN Directory, Paris.

ISSN - 2249-555X

Indian Journal of Applied Research

Journal for All Subjects

INDEX

Volume : 1 | Issue : 6 | March 2012 ISSN - 2249-555X

Research Paper Management

Factor Analysis and Business Research

* Dr. Kakali Majumdar

* Assistant Professor, College of Management, Shri Mata Vaishno Devi University, J&K

INTRODUCTION

Definition

Facto analysis in one of the main important tools to deal with large number of complex and diverse data sets. Basically it is an independent technique of multivariate data analysis, whose primary purpose is to define the underline structure among the variables in the analysis. Factor Analysis has widely been used in the field of mathematics, statistics, social sciences etc. No doubt, now a day, it is an inseparable part of business research.

By definition factor analysis refers to a variety of statistical techniques whose common objective is to represent large number of variables in to smaller number of hypothetical variables. In other word it can be said that the main purpose of factor analysis is the condensing of data.

There are many reasons why the business researcher uses factor analysis. The researcher may have a collection of good number of variables and would like to have some idea about what constructs might be used to explain the correlations among these variables. In brief factor analysis acts as a data reduction tool. It removes redundancy or duplication from a set of correlated variables and represents correlated variables with a smaller set of "derived" variables. Factors which are formed by this method, are relatively independent of one another.

History of factor analysis

The concept of factor analysis is originated from the study of Spearman (1904). He pioneered the use of factor analysis in the field of psychology. Spearman model with a single latent construct became less popular. In 1930's, L. L. Thurstone developed the common factor model with multiple latent variables,

Type of factor analysis

Two types of factor analyses, namely (i)exploratory factor analysis and (ii) confirmatory factor analysis are popularly used in practice.

Exploratory factor is the most common form of factor analysis. This analysis is used to reveal the number of factors and the variables that belong to specific factors. This type of Factor analysis is based on the priori assumption that any indicator may be associated with any factor. In exploratory factor analysis we do not have a pre-defined idea of the structure or how many dimensions are in a set of variables.

Assumption of factor analysis

Firstly the variables considered in factor analysis should be linearly related to each other. This can be checked from the scatter diagram. Secondly the variables must also be at least moderately correlated to each other, otherwise the number of factors will be almost the same as the number of original variables. This will result useless factor analysis.

Important Terminology

For analyzing factor analysis the following terminology must be taken into consideration

Loading: Loading is nothing but the factor-variable correlation. In the above equations α_{ij} s are the loadings.

Eigen Value: Sum of the square of the factor loadings is known as Eigen Value. How many principal components will be retained in the example is indicated by Eigon values. Principal component with eigon value more than one should be retained.

Communality: Another important terminology used in principal component method is the communality. It is the sum of the squared loading of each variable. The value of communality indicates the coverage of the variable by the representative factors.

Factor Rotation: If the initial factor loading matrix shows complex structure, the factor axis is rotated by an angle such that the factor loading are revised to have a simple structure. There are several method of factor rotation. Varimax rotation is the most widely used method of rotation that maximizes the variance of loadings within each factor.

Sample size selection

Different researchers have recommended different views regarding the size of the sample. Gorsuch (1983) recommended five subjects per item, with a minimum of 100 subjects, regardless of the number of items. Guilford (1954) argued that sample size should be at least 200. Cattell (1978) recommended three to six subjects per item, with a minimum of 250. Comrey and Lee (1992) suggested five state of sample size as 100= poor, 200 = fair, 300 = good, 500 = very good, 1,000 or more = excellent. The maximum no of sample size was recommended by Everitt (1975), minimum of 10 subjects per item. Anyways large sample size is desirable for factor analysis. Small samples may cause sampling error, which can manifest itself in factors that are specific to one data set. MacCallum et al (1999) suggest that increasing the sample size is one means of overcoming the problem of sampling error and with the large sample, factor analysis solutions become more stable and more reliably produce the factorial structure of the population.

MATHEMATICAL STRUCTURE

Its mathematical structure consists of multiple regression, product moment correlation, canonical analysis, partial correlation and analysis of variance. Thus it is theoretically capable of integrating many diverse items and obviously complex in nature.

The steps of factor analysis can be summarized in the following way:

Problem Formulation

⇓

Calculation of Correlation Matrix

⇓

Selection of method of Factor analysis

Determination of factor
⇓
Factor Rotation
⇓
Interpretation or result finalization

Even with the computational difficulties in recent years factor analysis become easily accessible to a wider circle of business research due to the development of user friendly computer packages like SPSS, SAS etc.

Example:
There are different methods of factor analysis like Principal Component Method, Centroid method, Maximum likelihood method etc, however, in the present paper we will be concentrating on the Principal Component Method developed by H. Hotelling. This method basically investigates whether a large number of variables of interest say, Y_1, Y_2,..........Y_n are linearly related to a smaller number of factors or components say F_1, F_2,..........F_k (k<n). Broadly data redundant is the primary objective of factor analysis.

Let us assume Y1, Y2……….Y6 are the six subjects of Management namely Organizational Behaviour (Y1), Economic Environment (Y2),Financial and Management Accounting (Y3), Computer Applications for Managers (Y4),Business Economics for managers(Y5),Quantitative Analysis (Y6).

Here the objective of the principal Component method is to find out the appropriate number of principal components (less than 6) using the following relationships

$$F_1 = \alpha_1 Y_1 + \alpha_2 Y_2 + ...+ \alpha_{1n} Y_n$$
$$F_2 = \alpha_2 Y_1 + \alpha_2 Y_2 + ...+ \alpha_{2n} Y_n$$

$$F_1 = \alpha_1 Y_1 + \alpha_2 Y_2 + ...+ \alpha_{1n} Y_n$$

Basically the grouping is based on some common characteristic. Like Financial and Management Accounting (Y3), Computer Applications for Managers (Y4), Quantitative Analysis (Y6) are scoring papers. The other papers are relatively less scoring. Is it possible to make such type of grouping based on the given information?

RUNNING FACTOR ANALYSIS IN SPSS:
Statistical Package for the Social Sciences (SPSS) is very popular, user friendly package for management as well as all other social science research. During 1968, Norman H. Nie, C. Hadlai (Tex) Hull and Dale H. Bent, developed SPSS based on the idea of using statistics to turn raw data into information essential to decision-making. The dialog box of it looks like figure-1.

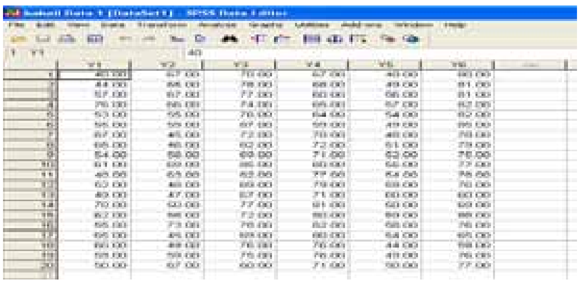

Figure -1

After entering the selected variables(in our case it is 6) in the dialog box of SPSS, come to the main dialog box of factor analysis by clicking **Analyze→ Data Reduction→ Factor Analysis** menu. The following dialog box will appear (Figure-2). From there select the variables to be included for factor analysis and transfer them to the dialog box named "Variables".

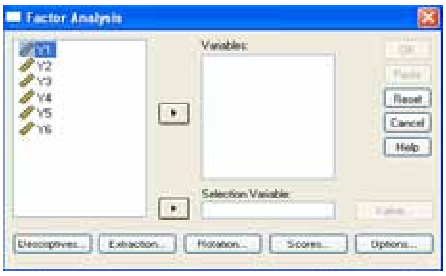

Figure -2

Now see the bottom of the Figure-2, 5 options are displayed **(Descriptive, Extraction, Rotation, Scores and Options).**

By clicking **Descriptive** the following dialog box (Figure-3) will appear.

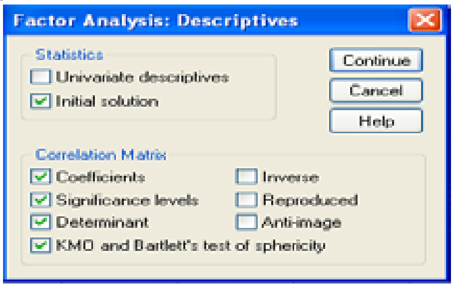

Figure-3

The coefficient option produces the R matrix. The significant option highlights the significance level of each correlation of the R matrix. The determinant option is the indicator of multicollinearity problem in the data set.

The second option of Figure -2, that is, Extraction will directly allow to select the desired method of Factor analysis (Figure-4). The principal component method is most widely used factor method in business research.

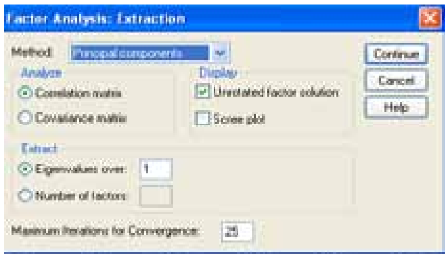

Figure-4

The Rotation, allows us to select the type of rotation we want to apply. Figure-5 indication the Varimax method has been considered.

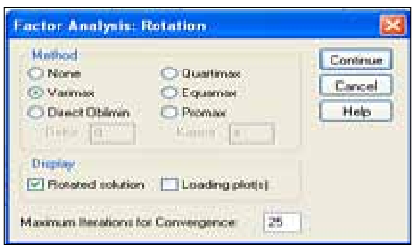

Figure-5

The Score option allows us to save factor scores for each subject in the data editor (see Figure-6). New columns will be automatically constructed in the original data sheet of SPSS need for further analysis like regression etc.

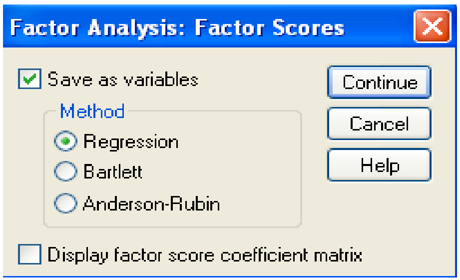

Figure-6

Clicking on Option (Figure-7) researcher can sort the data by size that will order the variables by their factor loadings.

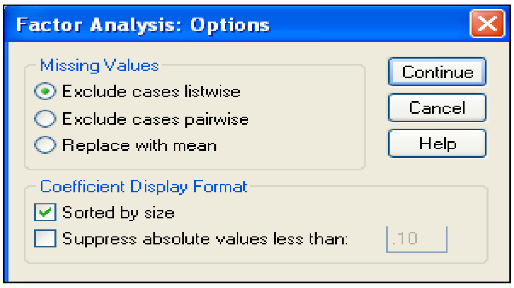

Figure-7

INTERPRETATION OF THE SPSS OUTPUT

Factor analysis begins with the construction of correlation matrix between data variables that are being studied. Consequently, R matrix or correlation matrix as shown in Table-1 will first appear on the output screen of SPSS. Out of two parts of the correlation matrix, the top half contains the correlation coefficient and the bottom half the significance level of the coefficients.

Table-1

Correlation Matri

		Y1	Y2	Y3	Y4	Y5	Y6
Correlation	Y1	1.000	-.325	.103	.082	.176	-.067
	Y2	-.325	1.000	.181	-.341	.022	.494
	Y3	.103	.181	1.000	-.310	-.049	.288
	Y4	.082	-.341	-.310	1.000	.153	-.394
	Y5	.176	.022	-.049	.153	1.000	.064
	Y6	-.067	.494	.288	-.394	.064	1.000
Sig. (1-tailed)	Y1		.081	.333	.365	.229	.390
	Y2	.081		.223	.071	.464	.013
	Y3	.333	.223		.092	.418	.110
	Y4	.365	.071	.092		.260	.043
	Y5	.229	.464	.418	.260		.394
	Y6	.390	.013	.110	.043	.394	

The second table shows the Kaiser-Meyer-Olkin (KMO) and Bartlett's test. The value of KMO statistic varies between 0 and 1. KMO value greater than 0.5 is desirable for factor analysis. KMO statistic with 0 value indicates the sum of the partial correlations is larger than the sum of simple correlation indicating factor analysis inappropriate. In our present example the KMO value 0.618 (Table-2) indicates the factor analysis is applicable.

Table-2

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.618
Bartlett's Test of Sphericity	Approx. Chi-Square	25.811
	df	15
	Sig.	.05

From Bartlett's test, the nature of the correlation matrix is tested. If there is any relationship between the variables then factor analysis has its meaning. If the correlation matrix is an identity matrix showing no relationship among the variables, factor analysis once again is not workable. The significant Bartlett test (Table-2) ensures relationship among the variables.

Table-3

Communalities

	Initial	Extraction
Y1	1.000	.793
Y2	1.000	.748
Y3	1.000	.646
Y4	1.000	.604
Y5	1.000	.883
Y6	1.000	.681

Table-3 is showing the communalities. The concept of communality has already been discussed. It is the sum of the squared loading of each variable. Two columns of Table-3 labeled initial and extraction indicate communalities before and after extraction respectively. Principal component analysis based on the assumption that all variance is common at the initial state indicating unit value of the communalities before extraction. Communalities in the next column say .79 indicates that 79% of the variance associated with variable one is common or shared variance. Again, after extraction only 19% of the original variable remains outside the explanation.

Table-4

Total Variance Explained

Component	Initial Eigen values			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.069	34.480	34.480	1.905	31.744	31.744	2.069	34.480	34.480
2	1.225	20.412	54.892	1.347	22.442	54.187	1.225	20.412	54.892
3	1.061	17.677	72.569	1.103	18.383	72.569	1.061	17.677	72.569
4	.654	10.900	83.470						
5	.557	9.290	92.760						
6	.434	7.240	100.000						

Table-4 lists the eigen values associated with each component. Result includes all the three varieties of eigen values i.e. before extraction, after extraction and after rotation. Before extraction there are 6 components and after extraction it is 3. The present factor analysis extracts 3 factor having eigen values greater than 1. From the values of the communalities it is possible to conclude that the present factor model explain highest proportion of variance of the variable Y3(communality 72.57%), while it is lowest for variable Y1.

Table-5 and Table-6 are the component matrices before and after extraction showing the loadings of each variable onto each factor.

Table-5
Component Matrix(a)

	Component		
	1	2	3
Y6	**.770**	.209	.212
Y2	**.764**	-.186	.361
Y4	-.718	-.066	.290
Y3	**.521**	.461	-.402
Y1	-.305	**.807**	-.221
Y5	-.115	.528	**.769**

Table-6
Rotated Component Matrix

	Component		
	1	2	3
Y4	-.735	.125	.220
Y3	**.728**	.319	-.115
Y6	**.727**	-.267	.284
Y1	.074	**.859**	.221
Y2	**.543**	-.639	.212
Y5	-.071	.105	**.931**

CONCLUSION

The principal component method in the present case, classifies the six set of variables into three common factors which associate variables with significant factor loadings and clarify the nature of mutual relation. Factor 1 represents variable Y2 (Economic Environment), Y3 (Financial and Management Accounting) and Y6 (Quantitative Analysis). Though theoretically, economic environment is less scoring paper, in the present case making group with Financial and Management Accounting and Quantitative Analysis. However, the loading of economic environment is less as compared to the other two subjects. The other subjects are not forming any group showing week correlation.

REFERENCES

Beri, G. C, (2000), Marketing Research, Tata McGraw-Hill Publishing Company Ltd. New Delhi. | Cattell, R. B. (1978), The Scientific Use of Factor Analysis. New York: Plenum. | Comrey, A. L. and Lee, H. B., (1992), A first course in factor analysis, Hillsdale, New Jersey: Erlbaum | Gorsuch, R. L. (1983), Factor Analysis (2nd. Ed). Hillsdale, NJ: Erlbaum. | Everitt, B. S., (1975), Multivariate analysis: the need for data and other problems, British Journal of Psychiatry, 126, 237-240. | Kim, J. O. & Charles Maeller (1978), Introduction to Factor Analysis, Sage University Paper. | Kim, J. O. & Charles Maeller(1978), Factor Analysis: Statistical Methods and Practical Issues, Beverly Hills, calif, Sage Publication. | Kothari, C.R (2004), Research Methodology: Methods and Techniques, New Age International (P) Ltd, New Delhi. | Pannerselvam, R. (2004), Research Methodology, Printice. Hall of India, New Delhi. | MacCallum, R. C., Widaman, K. F., Zhang, S. & Hong, S., (1999), Sample size in factor analysis, Psychological Methods, 4, 84-99. | Sperman, C (1904), General intelligence objectivel, Determined and Measured, American Journal of Psychology, 15, 201-294

Sara Publishing Academy
Indian Journal Of Applied Research
Journal for All Subjects

Editor,
Indian Journal Of Applied Research
8-A, Banans, Opp. SLU Girls College,
New Congres Bhavan, Paldi, Ahmedabad-380006.
Contact.: +91-9824097643 E-mail : editor@ijar.in

Printed at Unique Offset, Novatsing Rupam Estate, Opp. Abhay Estate, Tavdipura, Shahibaug, Ahmedabad