THE USE AND APPLICATION OF MULTIVARIATE ANALYSIS TECHNIQUES IN BIBLIOMETRIC AND SCIENTOMETRIC STUDIES

Farideh Osareh, Ph. D.
Department of Library & Information Sciencc,
School of Education & Psychology,
Shahid Chamran University, Ahvaz, I. R. of Iran
email: osareh_f@cua.ac.ir

Abstract - In Bibliometric and scientometric studies, three approaches of multivariate analysis, namely Factor Analysis, Cluster Analysis and Multidimensional Scaling are the most used methods. This article aims to review the use and applications of thcse techniques.

Keywords - Multivariate Analysis, Factor Analysis, Cluster Analysis, Multidimensional Scaling.

INTRODUCTION

In presenting the results and conveying the magnitude of relationships between pairs of units of analysis or variables, much use is madc of graphical displays in data analysis and especially in multivariate data analysis [1]. Generally, three approaches to multivariate analysis have been used in citation studies (including bibliometric and scientometric studies to display the inter-documentany relationships in the matrices of the most highly cited/co-cited documents) as follows: Cluster Analysis, Multidimensional Scaling and Factor Analysis. This article aims to review the use and applications of these approaches in bibliometric and scientometric studies.

MULTIVARIATE ANALYSIS

The branch of statistics dealing with procedures for summarizing, representing and analyzing multiple quantitative measurements of a number of individuals or objects is called 'Multivariate Analysis' [2]. Multivariate Analysis is described as the examination of several variables [3]. It is also described as the analysis of observations on several correlated random variables [4].

One aspect of this kind of analysis is the 'dimensionality reducing techniques', which include Cluster Analysis, Multidimensional Scaling and Factor Analysis. Since all these three methods intend to simplify a complex pattern of association among many variables, these techniques are called 'dimensionality reducing techniques'. In other words, simplification is done by projecting or representing an object in a higher-dimensional space, in a space of a smaller number of dimensions (usually two) [4].

Gcnerally, in bibliometric studies, data sets are large. For example, computerized databases are used to count publications, citations and patents [5]. On the other hand, the analysis of large sets of data is probably more complicated than smaller sets of data [6]. Therefore, the dimensionality reduction methods are relevant for bibliometric data analysis, since these methods are able to reduce data to fewer dimensions and also more interpretable groups.

CLUSTER ANALYSIS

One of the ancient and fundamental processes in science is classification. The facts and phenomena must be organized before we can understand them [7]. Children become familiar with classification, at quite an early age, by classifying the objects in the environment and by using nouns, in their languages, to apply to different classes of objects [8]. In other words, identifying the similarities and differences between the entities or objects that create the world can lead to a better understanding of the world. Objectively, grouping the entities together (i.e. classifying) on the basis of their similarities and differences is possible by Cluster Analysis [9]. Cluster Analysis (CA) is a multivariate procedure for detecting natural groupings in data [10]. Different kinds of procedures that could be used to create a classification are known as Cluster Analysis.

Experimentally, 'clusters' or groups of highly similar entities are formed by these procedures [8]. Although Cluster Analysis has been recognized in the current century, most of its literature has been provided during the past two decades [8]. The practice of Cluster Analysis was restated in computer terms in the 1950s to enable the investigator to escape from hand calculations [9].

Generally, Cluster Analysis is used to group objects, people, countries or other entities on the basis of shared characteristics [11]. The clustering of the literature is an effective tool for summarizing a large number of publications into a small number of categories. This clustering makes their interpretation easier [12].

In the next part of the article, the Agglomerative Hierarchical clustering method will be reviewed briefly.

AGGLOMERATIVE HIERARCHICAL CLUSTERING METHOD

In this method, clusters are formed by grouping cases into larger and larger clusters until all cases are members of a single cluster [13]. The first step of Agglomerative Hierarchical cluster analysis treats all cases as individual clusters, and its final step takes all the cases and merges them into one large group [8,14].

There are many criteria for deciding which cases or clusters should be combined at each step. All these criteria are based on a matrix of either dissimilarities or similarities between pairs of cases. A correlation or proximity matrix can be used as similarity measures among the cases [11]. Similarity, dissimilarity, correlation, overlap measure or any other variable for measuring similarities or differences between two objects of a single type are identified as a proximity [15]. The input matrixes to the procedures CLUSTER, ALSCAL and FACTOR are proximity matrix outputs[1] [13].

The Statistical Package for the Social Sciences (SPSS), for Windows, provides a clustering program that implements seven (i.e. six linkages plus one variant for Group-average method) different Agglomerative Hierarchical procedures[2] [13].

However, it has been claimed that from amongst at least twelve different linkage forms proposed, only four have become widely accepted [8], though some like [16] believe that there are only six linkages that are most widely used in practice. These linkages are as follows:

1. **Single linkage or nearest neighbor.** In this method, clusters are formed based on the similarities between the most similar pair of objects. For example, the first two cases combined are those that have the greatest degree of similarity [13].
2. **Complete linkage or furthest neighbor.** This method is the logical opposite of single linkage, i.e. the distance between two clusters is calculated as the distance between their two furthest (least similar) points; therefore, an object can be included into an existing cluster if it holds a certain similarity to all members of that cluster [8,13]. However, it is stated that in this method clusters themselves are still formed on the basis of the shortest 'distance' between clusters, just as in the single link method [4].
3. **Group-average method.** The distance between two clusters in this method is an average of the distances between all pairs of cases, where each member of the pair is from each of the clusters. For example, if cluster A forms a cluster with cluster B, the distance from cluster C to the new cluster AB is defined as the average of all distances between all points from C and all points from AB [13,16]. However, the average linkage within groups method is described as a variant of Group-average method that combines clusters, so that the average distance between all cases in the resulting cluster is as small as possible and, thus, the distance between two clusters is taken to be the average of the distances between all possible pairs of cases in the resulting cluster [13].
4. **Ward's method**. This method is slightly different from the three linkages discussed above. Although the reduction of the number of clusters continues, the distances between points are defined as differences between clusters rather than the similarities [4].
5. **Centroid method.** This method calculates the distance between two clusters as the distance of their mean values for each variable [16]. One of the disadvantages of the Centroid method is that if the sizes of the two groups, to be used, are very different, the centroid of the new group will be very close to that of the larger group and may remain within that group. The characteristic properties of the smaller group are, then, virtually lost [16].
6. **Median method.** In this method, equal weights are given to the two clusters being combined (i.e. assuming that the groups to be combined are of equal size) regardless of the number of cases in each group. This gives both groups an equal opportunity for representation in clusters that are merged [13,16].

All the above linkage procedures already exist in SPSS, from Release 6,0 for Windows to the latest versions.

Each clustering method produces the following graphic displays: a histogram-like icicle plot (vertical and horizontal), a tree-like dendrogram plot or both. It is believed that the most familiar expression of the results of these clustering methods is the dendrogram (tree diagram), which is a graphical display of the hierarchical structure implied by the similarity matrix and clustered by the linkage rule [8]. Similarly, it has been pointed out that the main outcome of a cluster analysis is a dendrogram, which is also called a tree diagram [13].

For example, Figure 1 is a dendrogram for the four-beer example used in the SPSS manual [13]. As indicated in Figure 1, Budweiser and Old Milwaukee were linked as the first two clusters by a line that is one unit from the origin. Next, Becks and Kirin are linked by a line that is 8 units from the origin. Likewise, when these two clusters are merged into a single cluster, the line that connects them is 25 units from the origin. This is the rescaled distance point, i.e. the more similar the beers, or the clusters, the sooner the joining will occur [14]. The dendrogram, therefore, shows not only which clusters are joined but also the distance (i.e. the degree of similarity) at which they are joined. In other words, Budweiser and Old Milwaukee, which are linked in a distance of one unit from the origin, are more similar than Becks and Kirin linked at a distance of 8 units from the origin (Figure 1).

Rescaled Distance Cluster Combine

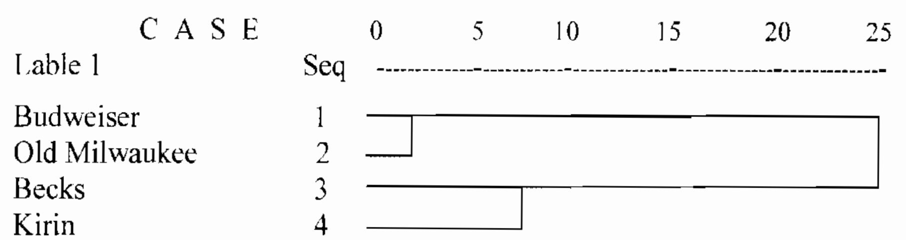

Figure 1: Dendrogram using complete linkage
Source: [13], p. 93

HOW MANY CLUSTERS TO KEEP OR WHERE TO CUT THE TREE?

Like the other techniques, Cluster Analysis presents the problem of how many factors, dimensions or clusters to keep. One rule of thumb for this is to choose a place where the cluster structure remains stable for a long distance. Some other possibilities are to look for cluster groupings that agree with the existing or expected structures, or to replicate the analysis on subsets of the data to see if the structures emerge consistently [17].

MULTIDIMENSIONAL SCALING (MDS)

MDS refers to a set of techniques which are used to represent the similarity of objects by points in space: the more two objects are similar, the smaller the distance between the two points will be and vice versa. Capturing the original data as much as possible in usually two or three dimensions (i.e., to 'reduce its space') is a major purpose of MDS [11].

MDS techniques use the matrices of proximity data (similarities, dissimilarities or distances) among any kinds of object as input. The main output is a spatial representation, consisting of a geometric configuration of points, as on a map. Each point is a representation for each object. The configuration reflects the 'hidden structure' in the data and it often makes the data much easier to comprehend [18]. Similarly, the 'hidden structure' within the matrices of proximity can be revealed by a MDS map in a small number of dimensions [19].

TYPES OF MDS

In general, there are two types of MDS:

- **Metric**
- **Non-metric**

Metric MDS makes the assumption that the input data is either ratio or interval, while the Non-metric model requires simply that the data be in the form of ranks. Therefore, the Non-metric model has far fewer restrictions than the Metric model and it is also less rigor. One technique to use if you are unsure whether your data is ordinal or interval is to try both Metric and Non-metric models. If the results are very close, the Metric model may be used [20].

Several different algorithms have been developed for multidimensional scaling based on Kruskal's work [21]. ALSCAL, is one algorithm which is used in most mapping programs of the citation studies [11,22,23]. It is an efficient and flexible program that performs a variety of MDS analyses on a large input matrix. The ALSCAL program minimizes the distortion with a statistic called 'S-stress' until improvement is less than 0.001[3][11,13].

The Stress value reported for each solution (usually Kruskal's Stress I or Stress II) and the proportion of variance explained (R Square in ALSCAL) are indicators of the overall 'goodness of fit' of that point configuration [11].

Essentially MDS works with only one matrix of proximity. Since a matrix is a two-way array, it is called a two-way MDS. When several proximity matrices are available for the same objects (perhaps one from each subject), the two-way MDS can analyze these data, but it does so by treating the differences among the matrices as due to random error. For this reason, a three-way MDS would be used if data are available in more than one proximity matrix [18]. The ALSCAL algorithm allows the analysis of an unlimited number of points or subjects in as many as six dimensions [24].

As in most citation analysis studies [11,23,25,26], this article used a two-dimensional solution of MDS (ALSCAL in SPSS) to assign relative locations to clusters based on the ranked order of their citation linkages [11,23,26]. The major reasons for selecting the two-dimensional solution have been mentioned in [18]. In this source, the authors focus on the ease of use of two-dimensional configurations versus those involving more dimensions. For example, a two-dimensional configuration is considered to be more useful than one involving three or more dimensions, when a MDS configuration is desired primarily as the foundation on which the clustering results can be displayed. The author of [14] confirms the above ideas and mentions some other reasons based on author co-citation analysis as follows:

1. Generally a high proportion of the variance (85% or more) is captured by a two dimensional map in the matrices of proximity and, consequently, a rich material for interpretation is provided [11].
2. The interpretation of MDS (author) co-cited maps is based on (author) point, placements and (author) cluster orientations along the horizontal and vertical axes.
3. A three dimensional solution generally adds little explanatory power. This is while it is more complex.
4. The examination of all possible map orientations, not simply those in the three or two-dimensional plots is required for the interpretation of a three-dimensional map.

In the mapping program, those variables sharing a high degree of similarity with many other variables, are placed near the center of the map. Others, not sharing a high degree of similarity, are placed near the edges of the map and represent more peripheral specialities [11].

As an example, in Table 1 used in the SPSS manual, a matrix of dissimilarity data entered using the SPSS Data Editor is presented [13]. Actually, these data are the flying mileages among 10 American cities. The cities are indicated as the 'objects' and the mileages as the 'dissimilarities'. For example, the distance between Atlanta and Chicago is 587 miles (shown in row 2, column 1), and that between Chicago and Denver is 920 miles (shown in row 3, column 2). Since all diagonal values represent the distance between a city and itself, they are all equal to 0. Each column of the matrix contains data for a different variable. The first column contains distances for Atlanta, the second column contains distances for Chicago and so on.

Table 1: Flying mileage between 10 American cities

	Atlanta	Chicago	Denver	Huston	LosAngeles	Miami	New York	Washington	Sanfrn	Seattle
1	0									
2	587	0								
3	1212	920	0							
4	701	940	878	0						
5	1936	1745	831	1374	0					
6	404	1168	1726	968	2330	0				
7	746	713	1831	1420	2451	1082	0			
8	2138	1858	949	1645	347	2594	2571	0		
9	2182	1737	1021	1891	950	2734	2408	678	0	
10	543	597	1494	1220	2300	923	206	2442	2329	0

Source: [13], p. 160

Figure 2 shows the MDS map based on these data. It shows the relative locations of 10 cities in the United States. The plot has 10 points, one for each city.

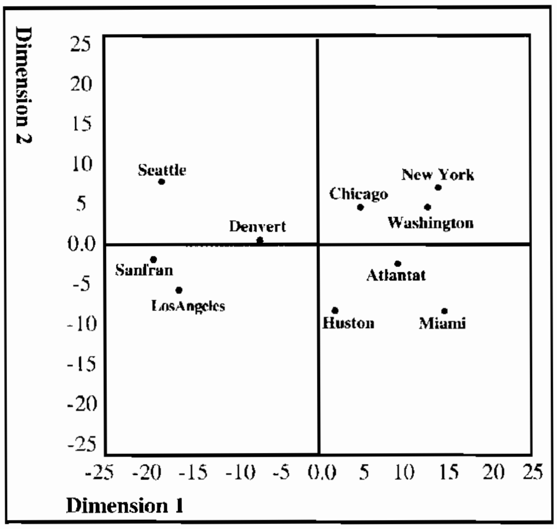

Figure 2: MDS plot of inte-rcity flying mileage
Source: [13], p. 160

Cities that are similar (have short flying mileages) are represented by points that are close together, e.g. New York and Washington are closer and are only 205 miles apart. However, cities that are dissimilar (have long flying mileages) are represented by points that are farther apart, e.g. Seattle and Miami are the furthest apart, there is a distance of 2734 miles between them.

FACTOR ANALYSIS

The term 'factor analysis' was first introduced by Thurstone in 1931. A relatively small number of factors, which can be used to represent relationships among sets of many interrelated variables, are identified as 'factor analysis' [13].

'Factor analysis' reduces a large number of observed variables to a fewer number and, in some ways, to more fundamental variables. It describes the results in the tests as functions of a few variables, which from certain standpoints are more 'fundamental' or convenient than the tests. This is usually done through an analysis of intercorrelation between the observed variables [27]. In other words, the main applications of factor analytic techniques are:

1- to reduce the number of variables

2- to detect structure in the relationships between variables, that is, to classify variables

Therefore, factor analysis is applied as a data reduction or structure detection method [28]. As an example in the SPSS manual [13], Table 2 shows 14 variables which are assigned only to 3 factors:

Table 2: Factor matrix

	Factor 1	Factor 2	Factor 3
Popstable	-.30247	.68597	.36451
Newscirc	-.67238	.28096	.49779
Femempld	.87461	.01131	.0863
Farmers	-.69659	.20002	-.40450
Retailing	.85161	.24264	.09351
Conoccrcl	.72503	.39394	.19896
Industzn	.84436	.29956	-.23775
Health	.38347	-.32718	-.63474
Childnegl	-.67430	-.12139	-.64221
Comoceffe	.63205	-.15540	.18706
Dwelgnem	.45886	-.73940	.24335
Migrrnpop	.07894	-.74371	.30110
Unemploy	-.78714	-.09777	.27134
Mentalit	-.30025	.45463	

Source: [13], p. 56

The variables within each Factor, i.e. Factor 1, 2, & 3 in Table 2, are related to each other according to their subject similarity Pcarson correlations coefficient, that is, a measure of similarity between two variables. Gencrally, in similarity measures large values indicate great similarity and small values indicatc little similarity among the variables [13]. The correlation coefficient can have valucs that range from -1.0 for a perfect inverse (negative) relationship, through 0 for no systematic association, up to +1.0 for a perfect direct (positive) relationship [29].

FACTOR EXTRACTION METHODS

Factor cxtraction is possible through Principal Components Analysis [13]. In this method, the obscrved variables are formed in the linear combinations, e.g. Factor 1, Factor 2, etc. (Tables 3-4). The combination that accounts for the largcst amount of the variance is thc first Principal Component (Factor 1) in the sample. Having thc second largest amount of variance and being uncorrelated with the first is the sccond Principal Component (Factor 2). Successive components explain progressivcly smaller accounts of the total sample variance, and all are uncorrelated with cach other [13]. For example, for this study, two factors are extracted by using Principal Component method: Factor 1 that accounts for thc largest portion of variance (75.7%) and Factor 2 that accounts for a smaller portion of variance (4.1%).

This study uses Principal Components Analysis, with an oblique factor rotation for extracting factors. Since factor analysis attempts to identify factors that arc substantively meaningful, i.e. summarized scts of closely related variables, the rotation phase of factor analysis attempts to transform the initial factor matrix into one that is easier to interpret. In other words, it will be possible to achieve factors in a simple structure by using a factor rotation [13]. There are two kinds of factor rotation: 'orthogonal' and 'oblique'. Thc orthogonal rotation creates uncorrelated (independent) factors, while an oblique factor rotation suggests whether factors arc independent by providing a matrix of factor intercorrelations [11,13]. Thc reasons for selecting an oblique rotation for this study are as follows:

- If the factors are uncorrelated (independent), it is unlikely that influences in nature are uncorrelated as well (e.g. the subject specializations they represent may not yet have been linked by citers)
- Even if the factors are uncorrelated in the population, they need not be so in the sample. Therefore, oblique rotations are more useful and have often been found to yield more significant factors [11,13]

Every variable loads on (contributes to) every factor, according to its factor loading[4]. The variables with a loading lower than + 0.3 are ignored[5]. In other words, factor analysis displays the variables in which the loadings represent the correlation of each variable with the factor. Moreover, in highly coherent fields, pointing to links between research specialties or other constructs, certain factors may have intercorrelations of +0.3 or above [11,13,30].

Authors may not appear more than once in a map or cluster but, they may contribute to more than one factor. Therefore, factor analysis may reveal particular facts about an author's 'breadth' where the other techniques do not [11].

NUMBER OF EXTRACTED FACTORS

How many factors are needed to represent the data? According to [13], to determine the number of factors one can choose one of the following two extraction criteria [11]:

1. **Eigenvalues over n:** The portion of variance accounted for each factor is called eigenvalue (e.g. for this study Factor 1 accounted for 75.7% of the variance with 31.8 eigenvalues and Factor 2 accounted for 4.1% of the variance with 1.7 eigenvalues). This gives some evidence of how important the factors are and helps resolve the question of how many factors exist in the data [11,13] (SPSS for Windows extracts factors with eigenvalues greater than one by default). The rationale introduced in [13] for selecting eigenvalues greater than one for the SPSS Factor Analysis procedure depends on a criterion that suggests including factors only with eigenvalues greater than one [13]. Since each variable has a variance of one, factors with a variance less than one are no better than a single variable. However, it has been asserted that in the SPSS Factor Analysis procedure, it is not always a good solution, although, this is the default criterion. It is possible to use a different eigenvalue rather than the SPSS default for factor analysis by choosing a number between 0 and the total number of variables in the analysis.
2. **Number of factors:** Extracting a user-specified number of factors is possible by entering a positive number regardless of their eigenvalues. (Generally this happens when the user knows how many clusters or factors exist in the data, e.g. 2, 3, etc.)

Like clusters, factors are extracted from the correlation matrix, but the cases are assigned to clusters based on their factor loadings [8].

The interpretation or definition of each new factor is based on those variables with high loadings. For example Figure 2, which is an example of a scree plot in the SPSS Manual, suggests that there are three factors in the data [13]. In other words, the initial statistics for principal components analysis, displayed in Table 3, show that there are only three factors with eigenvalues greater than 1: these three factors have accounted for nearly 72% of the variance observed. As can be seen, Table 3 contains six columns: the first two columns that provide information about the individual variables, are variables and communality. The communality is the proportion of variance explained by the common factors [13]. The communalities range between 0 to 1; 0 shows that the common factor indicates none of the variance and 1 indicates all the variance is explained by the common factor (Table 4). The other four columns describe the factors. The total variance explained by each factor in decreasing order is listed in the column labelled eigenvalue. For example, the linear combination formed by factor 3 has a variance of 2.00, which is 14.4% of the total variance of 14 (since each variable has a variance of 1, the 14 variables in this example have a total variance of 14) [13].

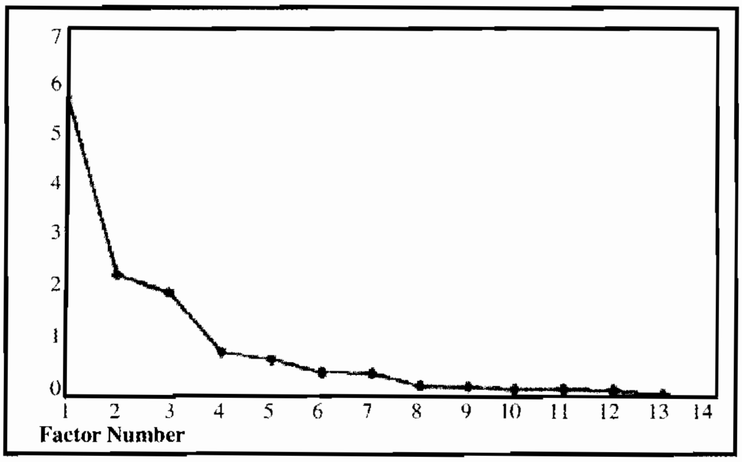

Figure 3: Factor scree plot

Source: [13], p. 55

Table 3: Factor initial statistics

Variable	Communality	•	Factor	Eigrnvalue	PCT of VAR	COM PCT
		•				
Popstable	1.00000	•	1	5.70658	40.8	40.8
Newscirc	1.00000	•	2	2.35543	16.8	57.6
Femempld	1.00000	•	3	2.00926	14.4	71.9
Farmers	1.00000	•	4	.89745	6.4	78.3
Retailing	1.00000	•	5	.75847	5.4	83.8
Conocercl	1.00000	•	6	.53520	3.8	87.6
Industzn	1.00000	•	7	.50886	3.6	91.2
Health	1.00000	•	8	.27607	2.0	93.2
Childnegl	1.00000	•	9	.24511	1.8	94.9
Comoceffe	1.00000	•	10	.20505	1.5	96.4
Dwelgnem	1.00000	•	11	.19123	1.4	97.8
Migrrnpop	1.00000	•	12	.16982	1.2	99.0
Unemploy	1.00000	•	13	.10202	.7	99.7
Mentalil	1.00000	•	14	.03946	.3	100.0

Source: [13], p. 54

Table 4, shows the final statistics. As displayed, the three factors with eigenvalue greater than 1 are picked up by default and the factors with eigenvalue lower than 1 are deleted automatically for this example.

Table 4: Factor final statistics

Variable	Communality	•	Factor	Eigrnvalue	PCT of VAR	COM PCT
		•				
Popstable	.69491	•	1	5.70658	40.8	40.8
Newscirc	.77882	•	2	2.35543	16.8	57.6
Femempld	.80696	•	3	2.00926	14.4	72.9
Farmers	.67503	•				
Retailing	.79253	•				
Conoccrcl	.72044	•				
Industzn	.85921	•				
Health	.65699	•				
Childnegl	.74921	•				
Comoceffe	.83607	•				
Dwelgnem	.79226	•				
Migrrnpop	.61855	•				
Unemploy	.71981	•				
Mentalil	.37047	•				

Source: [13], p. 57

Factors with eigenvalues lower than 1 are not important and arc ignored automatically; because, factors with a variance less than 1 are no better than a singlc variable [13]. Those columns labeled with 'Pct of var' and 'Cum pct' represent the percent and cumulative percentagc of variance.

CONCLUSIONS

Generally, in bibliometric and scientometric studies data sets are large. The analysis of large data is probably more complicated than the analysis of smaller sets of data. Therefore, the dimensionality reduction methods, which arc able to reduce data to a fewer dimensions and also more interpretable groups, seem relevant methods for bibliometric and scienctometric studies.

Especially, by using three methods of dimensionality reduction, namely Cluster Analysis, Multidimcnsional Scaling and Factor Analysis in bibliometric and scientometric studies, large and irrelcvant data will be reduced to a few significant and interpretable groups or dimensions. The irrelevant data will be ignored.

ENDNOTES

1. ALSCAL and FACTOR procedures are defined later.
2. Group-average method is onc of the six linkages that will be defined later.
3. S-stress (stress) is a measure of fit that is widely used in MDS. This measure of fit ranges from 1, worst possible fit, to 0 (perfect fit) [13]. In other words, the strcss measure is a criterion for determining the 'best fit' between the original input matrix 'distances' and the estimated distances in the chosen low-dimensional solution [11].
4. The factor loading here means the participation of each variable according to its

Pearson correlation coefficient.

5. The degree of association between a given empirical variable and a given factor can be indicated by a factor loading [3]. It is almost always +0.3 or above [11,13,30].

REFERENCES

[8] Aldenderfer, M. S. and Blashfield, R. K., *Cluster analysis*, Beverly Hills, CA, Sage Publications, 1984.

[3] Babbie, E., *The practice of social research*, 6th ed., Belmont, Wadsworth Publications, 1992.

[6] Bishop, Y. M. M., et al., *Discrete multivariate analysis: Theory and practice,* Cambridge, Mass, The MIT Press, 1975.

[5] Callon, M., et al., *Qualitative scientometrics,* In: Mapping the dynamics of science and technology: sociology of science in the real world. Ed. by M. Callon et al., Choundmills, The MacMillan Press, 1986.

[28] Copyright StatSoft, Inc. (1984-2003), *Principal Components and factor analysis, electronic textbook StatSoft,* available at: http://statsoftinc.com/textbook/stfacan.html.

[22] Cottrill, C. A., *A co-citation study of the scientific literature of two innovation research traditions: diffusion of innovations and technology transfer*, Ph. D. Dissertation, Ohio, The Union for Experimenting Colleges and Universities, 1987.

[30] Culnan, M. J., et al., *Intellectual structure of research in organizational behavior, 1972-1984: a co-citation analysis,* Journal of the American Society for Information Science, 41 (6), 1990, p.p. 453-458.

[1] Diday, E., et al., *Data analysis and informetric, III*, Proceedings of the Third International Symposium on data analysis and Informetrics, Amsterdam, North-Holland, 1984.

[4] Egghe, L. and Rousseau, R., *Introduction to informetrics: quantitative methods in library, documentation and information science,* Amsterdam, Elsevier Science Publishers, 1990.

[25] Everett, J. E. and Pecotich, A., *A combined loglinear/MDS model for mapping journals by citation analysis,* Journal of the American Society for Information Science, 42 (6), 1991, p.p. 405-413.

[16] Everitt, B., *Cluster analysis*, 3rd ed., London, Edward Arnold, 1993.

[27] Henrysson, S., *Applicability of factor analysis in the behavioral sciences: a methodological study*, Stockholm, Almquist & Wiksell, 1960.

[29] Hopkins, K. W., *Basic statistics for the behavioral sciences.* 2nd ed., Englewood Cliffs, Prentice-Hall Inc., 1987.

[15] Kruskal, J., *The relationship between multidimensional scaling and clustering,* In: Classification and Clustering: Proceedings of an advanced seminar conducted by the Mathematics Research Center, Wisconsin University at Madison, May 3-5, (1976) Ed. by J. Van Ryzin. New York, Academic Press, Inc., 1977, p.p. 15-44.

[18] Kruskal, J. B. and Wish, M., *Multidimensional scaling*, Beverly Hills, Sage Publications, 1978.

[21] Lane, J. G., *Mapping the cognitive structure of mathematics education using co-*

citation analysis, Ph. D. Dissertation, Kentucky, Kentucky University, 1984.

[14] McCain, K. W., *Core journals networks and co-citation maps: new bibliometric tools for serial research and management,* Library Quarterly, 61 (3), 1991, p.p. 311-336.

[19] McCain, K. W., *Longitudinal co-cited author mapping and intellectual structure: a test of congruence in two scientific literatures,* Ph. D. dissertation, Drexel, Drexel University, 1985.

[11] McCain, K. W., *Mapping authors in intellectual space: a technical overview,* Journal of the American Society for Information Science, 41 (6), 1990, p.p. 433-443.

[12] Miyamoto, S. and Nakayama, K., *A technique of two-stage clustering applied to environmental and civil engineering and related methods of citation analysis,* Journal of the American Society for Information Sciencc, 34 (3), 1983, p.p. 192-201.

[13] Norusis, M. J. *SPSS for Windows: professional statistics Release 5,* Chicago, SPSS Inc., 1992.

[23] Rice, R. E., et al., *Citation networks of communication journals, 1977-1985 cliques and positions, citations made and citations received,* Human Communication Research, 15 (1), 1988, p.p. 256-283.

[26] Sharma, L., *Mapping the citation links of journals in condensed matter physics,* In: Fifth International Conference of the International society for Scientometrics and Informetrics Proceedings, Ed. by M. E. D. Koenig & A. Bookstein, Medford, N. J., Learned information Inc., 1995, p.p. 515-524.

[7] Sokal, R., *Clustering and classification: background and current directions,* In: Classification and Clustering, Proceedings of an advanced seminal conducted by Mathematics Research Center, Wisconsin University at Madison, May 3-5, (1976) Ed. by J. Van Ryzin. New York, Academic Press, Inc., 1977, p.p. 1-15.

[2] Timm, N. H., *Multivariate analysis with applications in education and psychology*, Montery, California, Brooks/Cole Publications, 1975.

[17] Trochim, W. M. K., *Multivariate statistics: cluster analysis,* page available at: http://trochim.human.corncll.edu/tutorial/flynn/cluser.htm.

[20] Trochim, W. M. K., *Multivariate statistics: multidimensional scaling (MDS)*, page available at: http://trochim.human.cornel.edu/tutorial/msd.htm.

[9] Tryon, R. C. and Bailey, D. E., *Cluster analysis*, New York, McGraw-Hill Book Co., 1970.

[10] Wulder, M., *A practical guide to the use of selected multivariate statistics*, available at: http://www.pfc.forestr.ca/profiles/wulder/mvstats/introduction.html.

[24] Young, F. W., et al., *ALSCAL: a nonmetric multidimensional scaling program with several individual differences options,* Journal of Marketing Research, 15 (4), 1987, p.p. 612-615.