Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 6, Issue 2 (2013): 41-54

ORDERING MULTIVARIATE OBSERVATIONS

D.Vijaya Laxmi[1], V.V. Hara Gopal[2] and S.N.N.Pandit[3]
[1]Department of Statistics Kakatiya University, Warangal.
[2]Department of Statistics, Osmania University, Hyderabad.
[3]Center for Quantitative Methods, Osmania University, Hyderabad

(Received March 20, 2013)

Abstract

Intercropping data consists of multivariate observations of two or more crop-yields at different treatments (situations). Here, the problem is ordering these treatments with respect to the yields. At present land equivalent ratio (LER) is used to order the treatments and finding out at which treatment the LER has maximum value, considering that treatment is best one, to those intercrops. But, two or more crop-yields are not necessarily additive but may be interactive, then applying of LER could be misleading. In the present paper, we propose an alternative approach which brings out natural ordering if any exists, among the treatments.

Key Words: Ordering Multivariate Observations, Connectivity Approach, Principal Component Analysis, Inter cropping.

1. Introduction

Intercropping is the agricultural practice of cultivating two or more crops in the same space at the same time (Andrews & Kassam 1976). A practice often associated with sustainable agriculture and organic farming, intercropping is one form of poly culture, using companion planting principles. It is commonly used in tropical parts of the world and by various indigenous peoples (Altieri 1991), but in the mechanized agriculture of Europe, North America, and parts of Asia it is far less wide spread. Intercropping may benefit crop yield or control of some kind of pest, or may have other agronomic benefits.

In intercropping, there is often one main crop and one or more added crops, with the main crop being the one of primary importance because of economic or food production reasons. The two or more crops used in an intercrop may be from different species and different plant families, or they may simply be different varieties or cultivars of the same crop species, such as mixing two kinds of wheat seed in the same field.

The most common goal of intercropping is to produce a greater yield on a given piece of land by making use of resources that would otherwise not be utilized by a single crop.

2. Practices of intercropping cultivation

Finger millet is often intercropped with legumes such as peanuts, cowpeas, and pigeon peas, or other plants such as Niger seeds. Although statistics on individual millet species are confused, and are sometimes combined with sorghum. In this way the practice of intercropping are taken world over.

The Problem of ordering data is common enough to have attracted considerable attention of investigators. Generalizations of ANOVA and distribution free tests for the purpose have been in use for quite some time (Everitt and Dunn, 1999).

A natural extension of the problem arises when the characteristics of interest are not one but many in number. In such cases, it is rather than an exceptional situation that one has multivariate data which admits of the same (or, equivalently, exactly opposite) ordering between points with respect to each variable. One can, ofcourse, use concordance measures (like the concordance coefficient) to examine the overall consistency of ordering of the data on the basis of the each of the different variables separately. But even this admittedly 'weak' nonparametric procedure, will be of some use only to examine if they are arranging the data in the same order with respect to each of the variables. Consistency of ordering among the variables by possible rank reversal (i.e., the case of high negative correlation) will not be detectable here.

One natural way out, for solving this problem is to convert multivariate description of the data points into one variate description by a suitable transformation and solve this univariate ordering problem: Define a scoring function $g(x_1,x_2,\ldots..,x_k)$ of the k variables which makes sense in given context and 'reduce' the problem to one of univariate ordering. Though attractive, this approach is to be faulted: Ordering by g(x) naturally depends on the function g chosen for the purpose. Thus it is not a data determined ordering that we get, but a 'situation determined' one. When the situation (i.e., essentially the objective) changes, the function (and hence the ordering) will change. This approach, therefore, is essentially 'subjective' and does not reveal any structure inherent in the data.

In the context of intercropping, for instance, one may impute prices to the different output crops and use the net profit as the criterion for the ordering. This objective may be quite adequate if 'net money value' is the sole objective and stable. However, unless prices are stable enough, this criterion's utility is very questionable.

One such experimentation of intercropping data is being evaluated by way of ordering the multivariate observations with a new approach of Minaddition and found this approach to be meaningful and better than the so called existing method of Land equivalent ratio (LER).

3. Land equivalent ratio

One way to assess the benefits of intercropping is to measure productivity using the land equivalent ratio (LER). LER compares the yields from growing two or more crops together with yields from growing the same crops in monocultures or pure stands. Essentially the LER measures the effect of both beneficial and negative interactions between crops.

To calculate the LER, divide the intercrop yield of one crop by the yield of the pure stand and add that to the intercrop yield of the next crop divided by the yield of the pure stand and so on. The equation goes like this:

Intercrop1/Purecrop1 + Intercrop2 /Purecrop2 + etc. = LER

The resulting number is a ratio that indicates the amount of land needed to grow both crops together compared to the amount of land needed to grow pure stands of each. An LER greater than 1.0 usually shows that intercropping is advantageous and less than 1.0 shows a disadvantage. LER, ofcourse is perhaps a better criterion since it is apparently not situation-dependent. The disadvantage of LER method is that the two or more crop yields are not necessarily additive but could be interactive, and the impact of this fact on summing of LER is not clear. This disadvantage made us to investigate the new method of ordering the multivariate observations. Thus, we discuss the new methodology, first by considering the Principal component analysis and then deduce the new method.

4. Principal Component Analysis

Another approach which is data dependent is the Principal Component approach. It involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible.

PCA involves the calculation of the eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of component scores and loadings (Shaw, 2003). However, since the PC scores are highly scale dependent, there is a prima-facie case against their use except when the scales of measurement of the variables are pre fixed and unalterable. Also, when non-linearities are present, this approach fails. One component may be very inadequate to reveal nonlinear data structure as when for instance, the points form a closed loop. Search for alternative approaches to the problem of ordering multivariate data points, therefore, has to continue.

One such approach, found to be useful in the context of pattern recognition is presented below. This is the 'connectivity' approach to ordering of data-points. The concepts and theory of this approach is first presented. It is applied to some inter cropping data sets and its usefulness (or otherwise) will be briefly discussed. This approach does not force an ordering among data points but, if any meaningful ordering (not necessarily linear but, even, an open curved or close curved ordering) exists in the data, it reveals that ordering; if no such ordering exists that fact is also revealed by this analysis. Thus, if points lie on an open or closed curve, in a k-space that situation is revealed by this analysis. This approach has been found useful in Boundary alignment (or contouring) problems in Image processing, with considerable success (Pandit and Srinivas, 1989). In this illustrative case reports included in the sequel, we have only two dimensional data points and hence a visual display of the same is revealing.

The base for our procedure is the concept of Mahalanobis distances between data points. The k-vectors are utilized to obtain these distances in any meaningful way and the procedure take these distances as the base data.

Before proceeding with this approach, it is perhaps useful, for the sake of completeness to make a brief reference to the problem of multidimensional scaling as in PC concepts and Mahalanobis distance is of some relevance in the present context.

Mahalanobis distance is a distance measure introduced by Mahalanobis P.C. in 1936. It is based on correlations between variables by which different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample set to a known one.

Mahalanobis distance (or "generalized squared interpoint distance" for its squared value) can also be defined as a dissimilarity measure between two random vectors and of the same distribution with the covariance matrix *S*.

In order to use the Mahalanobis distance to classify a test point as belonging to one of N classes, one first estimates the covariance matrix of each class, usually based on samples known to belong to each class. Then, given a test sample, one computes the Mahalanobis distance to each class, and classifies the test point as belonging to that class for which the Mahalanobis distance is minimal.

5. The Connective Approach

Let $d_{i\,j}$ be the 'distances' between point pairs (i, j), i, j =1, 2, ---, N. A chain of points $\alpha_1, \alpha_2, \text{---}, \alpha_k$ is said to be constitute a path of k steps and of total distance $d_{i\,j}{}^{(k)} = d_{i\,\alpha 1} + \sum_{i=1}^{k-1} d\, \alpha_i\, \alpha_{i+1} + d\alpha_{kj}$ connecting i and j . A natural ordering among the points is obtained by choosing paths of smallest total distance (irrespective of step lengths) between point pairs. One can thus define shortest distances between point pairs and claim that the paths so obtained induce natural ordering among the points. These shortest distances and the paths to achieve the same can easily be obtained, by using the matrix operation of Minaddition (Pandit, 1962). Here, we first explain the method in detail with an illustration.

Minition: Let A and B be two matrices of real elements, then C = A . B is defined as the minition of A and B where $C_{ij} = \min(a_{ij}, b_{ij})$ e.g. $A = \begin{bmatrix} 3 & 2 & 5 & 4 \\ -1 & 10 & 2 & -5 \end{bmatrix}$,

$B = \begin{bmatrix} 5 & 10 & 4 & 6 \\ 3 & 2 & 7 & 6 \end{bmatrix}$ then $C = \begin{bmatrix} 3 & 2 & 4 & 4 \\ -1 & 2 & 2 & -5 \end{bmatrix}$

It is obvious that minition is analogous to the usual addition of matrices.

Minition (i) is defined only between matrices of the same order.

(ii) is commutative; $A.B = B.A$

(iii) is associative; $(A.B).C = A.(B.C)$

And (iv) obeys the transposition law; $(A.B)^T = A^T.B^T$

Minaddition: Let A and B be two conformable matrices. Then, minaddition is defined as $C = A \otimes B$ where $C_{ij} = \min_x (a_{ix} + b_{xj})$.

Let us consider $A = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 4 & -2 & -4 & 9 \end{bmatrix}$ and $B = \begin{bmatrix} 9 & 5 & 3 \\ 6 & -2 & 6 \\ 1 & -4 & 1 \\ 3 & 7 & 2 \end{bmatrix}$ then the minad product is $C = \begin{bmatrix} 4 & -1 & 4 \\ -3 & -8 & -3 \end{bmatrix}$

For instance, C (2, 3) = min (4+3, -2+6, -4+1, 9+2) = -3
Minaddition is obviously analogous to the usual matrix multiplication and it is defined only for ordered pairs of matrices in which the first matrix has the same number of columns as the second has the number of rows.

Minmaxion: Minmaxion is similar to minaddition except that, instead of taking the minima of sums, one takes the minima of the maxima in the pairs, and given by
C = A$\odot$B
i.e., $C_{ij} = \min(\max(a_{ix}, b_{xj}))$

Let $A = \begin{bmatrix} 3 & 5 & 2 \\ 1 & 3 & 6 \\ 9 & 3 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 & 2 & 5 \\ 7 & 3 & 4 & 2 \\ 7 & 5 & 3 & 9 \end{bmatrix}$ then minmax product of A and B is

$$D = \begin{bmatrix} 5 & 5 & 5 & 5 \\ 5 & 3 & 2 & 3 \\ 7 & 3 & 4 & 3 \end{bmatrix}$$

For instance, D (1, 1) = $\min[\max(3,5), \max(5,7), \max(2,7)] = 5$.
And it is defined only for the ordered pairs of matrices in which the first matrix has the same number of columns as the second has the number of rows.

A little modification of this concept of shortest distance between points leads to the proximity of ordering P items (points) on the basis of the degree of direct dissimilarity (distances) between them. Thus, for instance, let four items A, B, C, D have the mutual dissimilarities as given by the matrix entries d:

	A	*B*	*C*	*D*
A	0	10	3	9
B	10	0	20	7
C	3	20	0	30
D	9	7	30	0

Let 10 be a ‘threshold‘ score in that if two items have a dissimilarity of not more than 10, they are treated as ‘essentially equivalent‘. Thus on direct comparison C and D are most dissimilar with a score (distance) d =30 and are distinct from each other. But C and A have a dissimilarity of only 3 while A and D have a dissimilarity of only 9. Hence if one compares sequentially C and A, and A and D one can say that C, A, D are essentially equivalent, if one compares them in pairs, in that order. The definition of equivalence classes – through the chains of equivalent pairs of items is obviously useful in tracing an evolutionary chain among P items; (on the reasonable assumption that the evolutionary path is one with minimum dissimilarities along it).

In this context, we can also define a ‘connective distance‘ between the point pairs as the minimum score using which as a threshold we can say that the two points are equivalent.

For our problem of 4 items above, we get this connectivity distance matrix as

$$\begin{array}{ccccc} & A & B & C & D \\ A & 0 & 9 & 3 & 9 \\ B & 9 & 0 & 9 & 7 \\ C & 3 & 9 & 0 & 9 \\ D & 9 & 7 & 9 & 0 \end{array}$$

Thus with a threshold value of 6 , we conclude that A and C are equivalent; with threshold 8 , that A and C are an equivalent pair , and B and D are another equivalent pair while with a threshold of 10 all the four points are equivalent . Also, this threshold induces the following ordering C - A- D – B with distances 3, 9, 7 respectively.

Putting in a figurative way, let A B C D be four islands and a swimmer has the threshold capacity T of swimming at one stretch, a distance of 10 km. but not longer distance. Then he cannot reach directly C from B or D from C but he can reach them via the route indicated, with one or two intermediate landings. Thus, in respect of a threshold, items may be dissimilar (not mutually approachable) or similar (approachable) when viewed in a sequence, so that one goes, at each step for one point to another ‘similar’ point, along the ‘similarity path’. Naturalness of ordering the points along the similarity path, in many contexts, is obvious. One can compute the similarity paths connecting distances between all pairs of points by the matrix operation of minmaxion (Pandit & Srinivas 1989) $C = A \odot B$ where

$C_{ij} = \min_{x} (\max (a_{ix}, b_{xj}))$. The mathematics of operation is very similar to that of minaddition and (Pandit 1961, Das (1976), T.C.Reddy (1988)) presented elsewhere (ibid). Using this operation, one can thus define paths between any pair of points, ordering among them being essentially unique.

To recapitulate let a set of N data points be given. We first compute interpoint distances by employing suitable distance function like the Mahalanobis distance which is non-dimensional. Using the matrix minmaxion operation, we obtain the absolute connectivity level (the connective distance) between every pair of points. Choosing a threshold value we can partition (if so suggested by data) the given point set into clusters of connected sets, within each of which one can obtain an ordering among these points which is expected to be more meaningful than some of the other approaches like linear ordering through component scores.

This approach allows the data to speak for themselves rather than force a linear ordering when there may be none. It also allows one to recognize the possibility of even a closed loop structure in the data. This approach also allows developing indices of ordering and connectivity in data-sets which are more revealing than any other available approaches. However, as this aspect of the problem is not germane to the present objective of looking for meaningful ordering among multivariate data points, we shall not discuss it here anymore.

We shall now apply this method on three data sets from the field of intercropping. Data size is admittedly small, and each ‘point’ is an average vector of bivariate sample data. We shall not go into the reproducibility and other conventional inferential aspects; we shall rather illustrate our approach by using these data for analysis and comparing the results with those obtainable by PCA.

Since the data are bivariate only, graphical display itself gives some useful information which can be used to examine the validity and sensibility of the results from these different approaches.

For each data set of the dispersion matrix, the first principal component scores, and line diagrams for ordering the points as per these score differences then the Mahalanobis distance and connective matrices, followed by the tree diagram of the data as revealed by the connective distance matrix are presented.

6. Computational procedure to obtain tree

From the distance matrix D of given N data points, by applying the Min-Maxion operation on the distance iteratively, for some finite integer n, D^n=D^{n+1}. That D^n is the connectivity matrix C for given distance matrix D. The tree diagram can be obtained by defining the matrix P as follows,

P_{ij}=1 if $d_{ij} = c_{ij}$
=0, otherwise

If P_{ij} = 1 means i and j are connected directly. Using this concept the tree can be constructed.

We applied this method on three data sets and these data sets are collected from an agricultural experiment station in Andhra Pradesh during the year July 2006.

Data set – I

Yields for different dates of transplantation of Finger millet intercropped with Pigeon pea.

Treatment number	Treatment	yield(q\ha)	
		pigeon pea	finger millet
1	Sole pigeon pea (90x20cm^2) sown on July 4	15.6	----
2	Sole pigeon pea (75x25cm^2) sown on July 4	14.8	----
	Sole Finger millet		
3	Transplanted 20 days after pigeon pea sowing	----	13.5
4	Transplanted 30 days after pigeon pea sowing	----	10.9
5	Transplanted 40 days after pigeon pea sowing	----	6.9
	Intercropping		
6	Pigeon pea (90x20) + finger millet (20 days)	8.3	5.3
7	Pigeon pea (90x20) + finger millet (30 days)	13.4	0.9
8	Pigeon pea (90x20) + finger millet (40 days)	14.5	2.6
9	Pigeon pea (75x25) + finger millet (20 days)	11.0	4.7
10	Pigeon pea (75x25) + finger millet (30 days)	13.1	0.6
11	Pigeon pea (75x25) + finger millet (40 days)	14.9	0.6

Dispersion matrix

$$\begin{bmatrix} 42.19 & -27.94 \\ -27.94 & 21.48 \end{bmatrix}$$

% of variation explains by the first PC: 96

First PC scores

1	2	3	4	5	6	7	8	9	10	11
0.93	0.85	-1.68	-1.49	-1.20	-0.22	0.64	0.63	0.11	0.64	0.81

The order of the treatments obtained by the first PC scores:

1 , 2 , 11 , 7 , 10 , 8 , 9 , 6 , 5 , 4 , 3

In the distance and connectivity matrices, the lower diagonal values are identical to those of upper diagonal values.

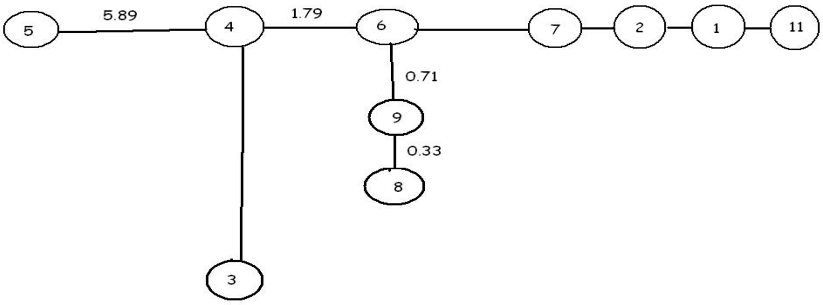
Distance Matrix

$$\begin{bmatrix} 0.000 & 0.120 & 10.054 & 6.465 & 10.675 & 1.470 & 0.240 & 1.323 & 1.561 & 0.573 & 0.020 \\ & 0.000 & 10.762 & 6.158 & 8.806 & 1.468 & 0.051 & 2.128 & 2.135 & 0.177 & 0.164 \\ & & 0.000 & 2.492 & 16.059 & 4.494 & 9.807 & 6.105 & 4.003 & 11.062 & 9.085 \\ & & & 0.000 & 5.898 & 1.800 & 5.151 & 6.103 & 3.586 & 5.450 & 5.858 \\ & & & & 0.000 & 7.388 & 7.720 & 15.834 & 12.677 & 6.549 & 10.471 \\ & & & & & 0.000 & 1.064 & 1.729 & 0.710 & 1.455 & 1.076 \\ & & & & & & 0.000 & 2.205 & 1.953 & 0.094 & 0.236 \\ & & & & & & & 0.000 & 0.337 & 3.209 & 1.114 \\ & & & & & & & & 0.000 & 2.822 & 1.046 \\ & & & & & & & & & 0.000 & 0.608 \\ & & & & & & & & & & 0.000 \end{bmatrix}$$

Connectivity Matrix

$$\begin{bmatrix}
0.000 & 0.120 & 2.492 & 1.800 & 5.898 & 1.064 & 0.120 & 1.064 & 1.064 & 0.120 & 0.020 \\
 & 0.000 & 2.492 & 1.800 & 5.898 & 1.064 & 0.051 & 1.064 & 1.064 & 0.094 & 0.120 \\
 & & 0.000 & 2.492 & 5.898 & 2.492 & 2.492 & 2.492 & 2.492 & 2.492 & 2.492 \\
 & & & 0.000 & 5.898 & 1.800 & 1.800 & 1.800 & 1.800 & 1.800 & 1.800 \\
 & & & & 0.000 & 5.898 & 5.898 & 5.898 & 5.898 & 5.898 & 5.898 \\
 & & & & & 0.000 & 1.064 & 0.710 & 0.710 & 1.064 & 1.064 \\
 & & & & & & 0.000 & 1.064 & 1.064 & 0.094 & 0.020 \\
 & & & & & & & 0.000 & 0.337 & 1.064 & 1.064 \\
 & & & & & & & & 0.000 & 1.064 & 1.064 \\
 & & & & & & & & & 0.000 & 0.020 \\
 & & & & & & & & & & 0.000
\end{bmatrix}$$

Tree Diagram



Data Set II

Yield values for various combinations of plant densities for Maize + Pigeon pea system

Treatment number	Treatment	Yield (q/ha)	
		Maize	Pigeon Pea
1	Sole maize (60cm)	26.6	-----
2	Sole maize (75cm)	22.0	-----
3	Sole pigeon pea (60cm)	-----	27.0
4	Sole pigeon pea (75cm)	-----	23.0
	Maize + pigeon pea		
5	100% + 100% (60cm)	22.9	9.3
6	100% + 100% (75cm)	18.7	19.4
7	50% + 50% (60cm)	15.9	22.7
8	50% + 50% (75cm)	11.8	17.3
9	Paired row (30/120cm)	12.8	13.7
10	Paired row (45/105cm)	22.5	7.6

Dispersion Matrix

$$\begin{bmatrix} 86.75 & -73.06 \\ -73.06 & 91.20 \end{bmatrix}$$

% of variation explains by the first PC: 91
First PC scores:

1	2	3	4	5	6	7	8	9	10
-1.41	-1.15	1.57	1.34	-0.67	0.11	0.45	0.37	0.12	-0.75

The order obtained by the first PC scores:
3, 4, 7, 8, 9, 6, 5, 2, 1

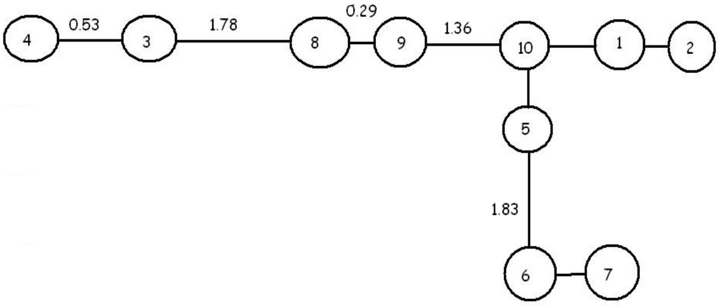
Distance Matrix

$$\begin{bmatrix} 0.000 & 0.833 & 9.853 & 9.076 & 1.607 & 6.884 & 8.484 & 3.681 & 2.601 & 0.859 \\ & 0.000 & 8.886 & 6.947 & 3.799 & 10.486 & 12.029 & 4.173 & 2.411 & 2.413 \\ & & 0.000 & 0.599 & 6.811 & 6.967 & 6.334 & 1.785 & 2.336 & 6.491 \\ & & & 0.000 & 7.886 & 10.008 & 9.658 & 2.456 & 2.181 & 6.956 \\ & & & & 0.000 & 1.839 & 2.737 & 1.646 & 1.938 & 0.157 \\ & & & & & 0.000 & 0.134 & 2.954 & 4.709 & 2.955 \\ & & & & & & 0.000 & 3.151 & 5.173 & 3.967 \\ & & & & & & & 0.000 & 0.298 & 1.484 \\ & & & & & & & & 0.000 & 1.386 \\ & & & & & & & & & 0.000 \end{bmatrix}$$

Connectivity Matrix

$$\begin{bmatrix} 0.000 & 0.833 & 1.785 & 1.785 & 0.859 & 1.839 & 1.839 & 1.366 & 1.366 & 0.859 \\ & 0.000 & 1.785 & 1.785 & 0.859 & 1.839 & 1.839 & 1.366 & 1.366 & 0.859 \\ & & 0.000 & 0.599 & 1.785 & 1.839 & 1.839 & 1.785 & 1.785 & 1.785 \\ & & & 0.000 & 1.785 & 1.839 & 1.839 & 1.785 & 1.785 & 1.785 \\ & & & & 0.000 & 1.839 & 1.839 & 1.366 & 1.366 & 0.157 \\ & & & & & 0.000 & 0.134 & 1.839 & 1.839 & 1.839 \\ & & & & & & 0.000 & 1.839 & 1.839 & 1.839 \\ & & & & & & & 0.000 & 0.298 & 1.366 \\ & & & & & & & & 0.000 & 1.366 \\ & & & & & & & & & 0.000 \end{bmatrix}$$

Tree Diagram



Data set III

Yield values due to N levels for Sorghum + Pigeon pea system

Treatment number	Treatment N(kg/ha)	Yield(q/ha)	
		Sorghum	Pigeon pea
1	15	12.4	14.8
2	30	16.4	16.4
3	45	19.3	18.8
4	60	20.1	17.9
5	75	22.9	17.5
6	90	24.6	17.0
7	105	26.9	16.4
8	120	28.7	15.0

Dispersion matrix:

$$\begin{bmatrix} 29.73 & 0.10 \\ 0.10 & 1.89 \end{bmatrix}$$

% of variation explained by first PC: 94

First PC scores:

1	2	3	4	5	6	7	8
-1.65	-0.91	-0.38	-0.23	0.27	0.58	1.01	1.33

The ordering of treatments obtained by the first PC scores:
8, 7, 6, 5, 4, 3, 2, 1

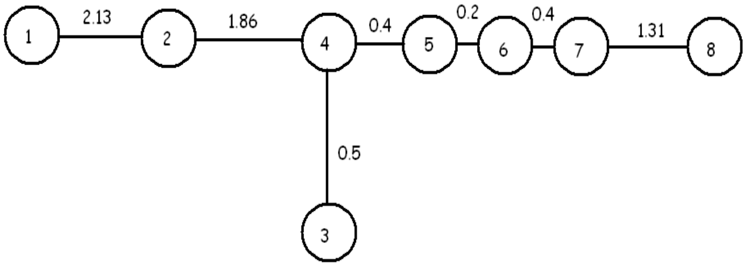
Distance Matrix

$$\begin{bmatrix} 0.000 & 2.133 & 11.372 & 7.980 & 8.520 & 8.531 & 9.531 & 10.225 \\ & 0.000 & 3.771 & 1.861 & 2.325 & 2.782 & 4.239 & 7.072 \\ & & 0.000 & 0.517 & 1.538 & 3.076 & 5.775 & 12.265 \\ & & & 0.000 & 0.403 & 1.285 & 3.179 & 8.025 \\ & & & & 0.000 & 0.266 & 1.364 & 5.127 \\ & & & & & 0.000 & 0.427 & 3.096 \\ & & & & & & 0.000 & 1.318 \\ & & & & & & & 0.000 \end{bmatrix}$$

Connectivity Matrix

$$\begin{bmatrix} 0.000 & 2.133 & 2.133 & 2.133 & 2.133 & 2.133 & 2.133 & 2.133 \\ & 0.000 & 1.861 & 1.861 & 1.861 & 1.861 & 1.861 & 1.861 \\ & & 0.000 & 0.517 & 0.517 & 0.517 & 0.517 & 1.318 \\ & & & 0.000 & 0.403 & 0.403 & 0.427 & 1.318 \\ & & & & 0.000 & 0.266 & 0.427 & 1.318 \\ & & & & & 0.000 & 0.427 & 1.318 \\ & & & & & & 0.000 & 1.318 \\ & & & & & & & 0.000 \end{bmatrix}$$

Tree diagram



The first data set consists of yields for different dates of transplantation of finger millet intercrop with pigeon pea. The finger millet was transplanted at three different dates in standing crop of pigeon pea sown on with two geometries.

By PCA it is found that the first principal component is enough to summarize the whole data since it explains 96% of the total variation. By the ordering of the first principal component scores, it is observed that at the two ends the sole crop treatments and in between the rest of the intercrop treatments are same.

The connectivity approach reveals neither perfect linear ordering nor circular ordering exists among the treatments. The sole crop treatments stood at two ends of the dendogram and they are too far from the remaining treatments 6, 8, 9. The treatments 6, 8 and 9 are formed as one more branch to the tree.

The second data set consists of Maize and pigeon pea intercrops data. Since the first principal component explains 94% of total variation, and the first principal component scores are enough to order the treatments. As in the first data set the two sole crops treatments are same at two ends of that order sequence.

By connectivity approach, it reveals neither linear ordering nor circular ordering exists among the treatments. In this also the sole crop treatments at the two ends and the rest of the treatments in between the sole crop treatments came in tree. The treatments 10, 5, 6, 7 formed one more branch to the tree.

The third data set consists of intercrop yields of Sorghum and Pigeon pea. The first principal component of the data explains 94% of total variation and the treatments are ordered according to the magnitudes of first principal component scores.

By the connectivity approach it reveals that there exists linear ordering and not circular ordering. In the tree structure the treatment 3 formed a branch but it is not too far to the stem of the tree. Except this treatment 3, the order of the remaining treatments is same as the ordering obtained by first PC scores.

7. Conclusion

In all the data sets seen we find that by PCA it was observed that the first PC is enough to summarize the whole data, since it explains 96% of the variation for the first data set, 94% of the variation for second and third data sets. Whereas the connectivity approach reveals neither perfect linear ordering nor circular ordering exists among the treatments. Thus, PCA in all the data sets shows a linearing ordering.

Therefore, in summary the ordering obtained by PC scores of the first and second data sets (no sole crop in the third data set), the sole crops came at the two ends in the order sequence. Connectivity approach reveals that no perfect linear ordering exists in the first two data sets. Third data set analysis reveals the following: Except the treatment 3, the order of the remaining treatments obtained by Connectivity approach is same as the order obtained by the first PC scores.

References

1. Andrews, D. J. and Kassam, A.H. (1976). The importance of multiple cropping in increasing world food supplies, pp.1-10 in R. I. Papendick, A. Sanchez, and G.B. Triplet (Eds.), Multiple cropping, ASA Special Publication 27, American Society of Agronomy, Madison, WI.

2. Altieri, M. A. (1991). The Ecology and management of insect pests in Traditional Agroecosystems, In Ethnobiology: Implications and Applications, proceedings of the first international congress of Ethnobiology, p.131-143, Brazil.
3. Everitt, B.S. and Dunn, G. (1999). Applied Multivariate data analysis – Edward Arnold- A division of hodder and Stoughton-London.
4. Shaw, P. J. A. (2003). Multivariate Statistics for the Environmental Sciences, London: Hodder-Arnold, ISBN o-3408-0763-6.
5. Pandit, S.N.N. and Srinivas, K. (1989). A sampling procedure for Image processing, presented at International Conference on Recent Developments in Statistical Data Analysis and Inference, August, Switzerland.
6. Pandit, S.N.N. (1961). A new matrix calculus, Journal of the Society for industrial and applied mathematics, 9, p. 623-639.
7. Pandit, S.N.N. (1962). Minaddition and an algorithm to find the most reliable paths in a Network, IRE transactions on circuit theory, ct.-9, p. 190-191.
8. Shila Das. (1976). Routing and Allied Combinatorial Programming Problems, Ph.D. Thesis, Dibrugarh University, Assam.
9. Reddy, T.C. (1988). on routing and related problems, M.Phil. Dissertation, University of Hyderabad, Hyderabad.