

Ordering of Hierarchical Classifications

Dimitar L.Vandev*, Yanka G.Tsvetanova †

Abstract: A new version of the single-link hierarchical clustering algorithm is presented. It produces a dendrogram which gives better graphical presentation of the proximity between the observed objects than the standard algorithms. A definition of a *perfect* chain is proposed. This kind of chains are useful for describing interesting properties of the algorithm. Some sufficient conditions for the shortest trajectory connecting all objects are included. Another useful property of the main idea is that the dendrogram produced by any clustering algorithm may be rearranged to get better interpretation.

Keywords: dendrogram, ultrametrics, order, compatible ordering, chain, perfect chain, minimum spanning tree

1 Introduction

Hierarchical cluster analysis attempts to group the objects of an observed set, on the basis of similarity or dissimilarity between them, into mutually exclusive subsets (clusters) which consist of close objects. These clusters may be grouped into larger sets and so on, until all objects are eventually united in one cluster. The higher the level of aggregation is, the less similar are the objects of the respective cluster.

The graphical representation of hierarchical classification is a tree-like diagram called *dendrogram*. The observed objects in it are terminal nodes of the tree. The sequence of joining the clusters is visualized by fusion of two nodes into a parent node and so on until all nodes (clusters) are united into a single node at the top of the diagram. Usually a scale is incorporated into the dendrogram to indicate the dissimilarity level (aggregation distance) at which the two nearest clusters are supposed to join.

The dendrogram with n nodes can be considered as one of 2^{n-1} binary trees. In Fig. 1 below there are four dendrograms presenting one and the same hierarchical classification. Each one of them may be obtained from another by rotating some vertical branches. They are indistinguishable with regard to nesting and amalgamation levels, and differ only in the arrangement of the terminal nodes and may therefore present a slightly different visual impression of the structure to the viewer.

¹Postal address: Laboratory of Computer Stochastics, Institute of Mathematics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria. Supported by National Foundation of Scientific Investigations, grant I-19/91.

²Postal address: Dep.of Math.and Phys., Thracian University, 6000 Stara Zagora, Stud.Grad, Bulgaria.

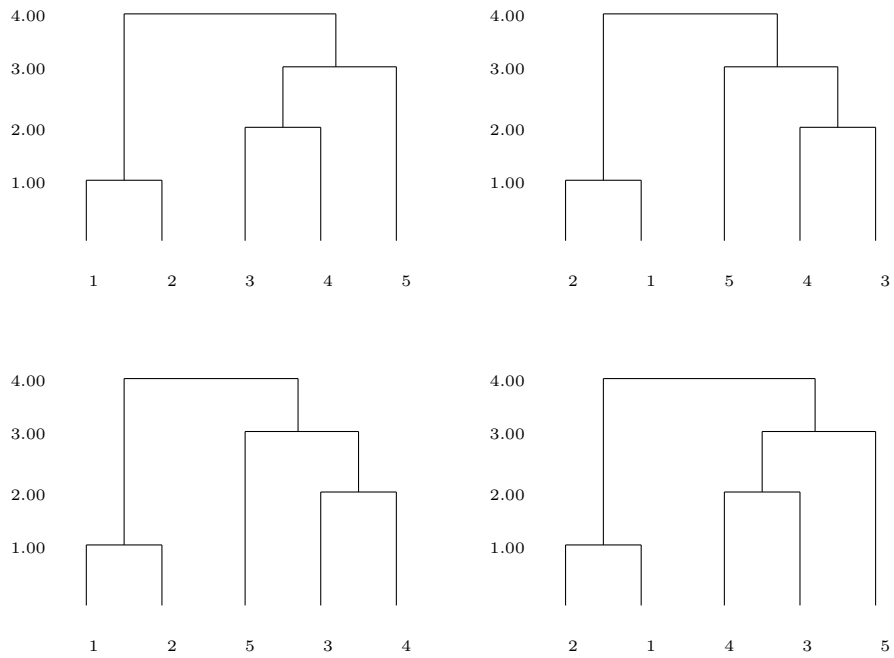


Fig. 1

There is a wide variety of methods for hierarchical cluster analysis. Different clustering methods imply different definitions of the dissimilarities (distances) between clusters and it is reasonable to give different hierarchical classifications for the same data sets. In all standard hierarchical methods there are no special rules for ordering of terminal nodes in the dendrogram as clusters are considered as usual sets of objects and to join together two clusters means to unite two sets.

The idea of arranging the nodes of the dendrogram in order to optimize some kind of criterion is not quite new. The problem has two sides: From among all possible representations of one dendrogram obtained via given hierarchical clustering method to be chosen such ones whose terminal node arrangement reflects in the best way the relationships between them (the similarity or dissimilarity); The other side of the problem of optimization arises as the simultaneous seeking of the hierarchical classification and the representation of this classification from among a set of possible classifications and their representations which optimize the criterion. This criterion may be the dissimilarity matrix reflecting the likeness or unlikeness among the objects or some external criterion (variable, order, another hierarchy) (Brossier ,1984).

Theoretically the solution of the optimal fitting problem is always possible by enumerating the 2^{n-1} trees and selecting the one tree whose terminal node ordering correlates most highly with the criterion. Such approach is impractical, however, because the value of 2^{n-1} is very large, even for not very large values of n . Therefore, it is of interest to find an approach for achieving an optimal ordering in a smaller number of steps.

Gruvaeus and Wainer (1972) present an algorithm applying series of tests for local orienting the nodes so that object displayed on the left and right edges of each cluster are adjacent to those objects outside the cluster to which they are most similar. Their procedure of ascendant hierarchy aims to obtain ordering of the terminal nodes in a chain with minimal length.

In some consistent papers Brossier (1980,1982,1983,1984) presents a theoretical base of the ordering problem and some algorithms for ordering of binary hierarchies. Degerman (1982) describes a procedure for orienting the nodes of a binary tree to maximize the Kendall rank order correlation τ between node order and a given external criterion.

In this paper we present an algorithm of ordering the terminal nodes. Our idea is to obtain an unique dendrogram using a specific way for linking the clusters in order to improve the visual representation of the hierarchical structure and to make them more realistic. This algorithm is appropriate for rearranging the nodes of the dendrograms obtained by the standard clustering methods.

2 Notations

Let X is a set of n objects $\{o_1, o_2, \dots, o_n\}$. and $\delta(.,.)$ is a positive real-valued symmetric function on $X \times X$, whose values reflect the relative closeness (similarity) or distance (dissimilarity) of objects to one another. Obviously, the smaller the similarity value or larger the dissimilarity one is, the further apart the associated objects are. Frequently it is more natural to consider dissimilarities. Of course, we can always convert a similarity matrix to a dissimilarity one or vice versa by some non increasing transformation. For convenience we will use the dissimilarity coefficients $\delta_{ij} = \delta(o_i, o_j)$, $o_i, o_j \in X$ and let D is the dissimilarity matrix. S denotes the set of all subsets of X and consists of 2^n elements. Benzecri (1965) gave the following definition of the indexed hierarchy.

Definition 2.1 *Indexed hierarchy on X is a pair (H, g) where H is a collection of elements of S and g is a function from H to R_+ which satisfy:*

- H1.** $X \in H$;
- H2.** $\forall a \in X, \{a\} \in H$;
- H3.** $C \cap C' = \{C, C', \emptyset\}, \forall C, C' \in H$
- I1.** $C' \subset C \implies g(C') < g(C), \forall C, C' \in H$
- I2.** $g(C) = 0 \iff C = \{a\}$ for some $a \in X$

The aim of the hierarchical clustering techniques is to construct an indexed hierarchy by use of dissimilarity coefficients, satisfying above the properties. The weakest partitioning contains n classes (clusters) and each one of them consists of one object. The strongest clustering contains all objects united into a single class.

For each indexed hierarchy (H, g) we define one ultrametric matrix U . The elements of this matrix have the following properties.

$$\forall a, b \in X \quad u(a, b) = \min\{g(C) : a \in C, b \in C, C \in H\}.$$

The ultrametric matrix is a dissimilarity matrix which satisfies the following property:

$$\forall a, b, c \in X \quad \text{we have } u(a, c) \leq \max\{u(a, b), u(b, c)\}.$$

It may be shown that U is an ultrametric iff all the triangles a, b, c , $\forall a, b, c \in X$ are isosceles

with a base smaller than the sides. There exist bijection between indexed hierarchies and ultrametrics (Benzecri, 1973).

In our algorithm we consider every cluster as an ordered set in a sense of the relative position of the objects each to another and call it *chain*. In this case the union of two clusters is not uniquely determined.

Definition 2.2 *Chain is a set with defined binary order relation θ between its elements with the properties:*

- (i) $a\theta a$ (*reflexive*)
- (ii) $a\theta b$ and $b\theta a \Rightarrow a = b$ (*anti-symmetric*)
- (iii) $a\theta b$ and $b\theta c \Rightarrow a\theta c$ (*transitive*)

$$\forall a, b, c \in X.$$

According to our considerations two chains consisting of the same objects ordered in a inverse way are equivalent (e.g. the chains $\{o_1, o_2, o_3, o_4\}$ and $\{o_4, o_3, o_2, o_1\}$).

Definition 2.3 *Let D is the matrix of dissimilarity coefficients between the objects of X and O is a chain of objects in X . O is compatible with D when*

$$a\theta b\theta c \iff \delta(a, c) \geq \max\{\delta(a, b), \delta(b, c)\}$$

$$\forall a, b, c \in O.$$

If O contains all elements of X , O will be compatible with the dissimilarity matrix D if all of her subchains are compatible with D . The aim of the ordering algorithms is to be found such arrangement of all elements in X compatible with the distance matrix D [3]. We should note that in general not for every dissimilarity matrix is possible to be found a compatible order. Although, if such compatible order exists and all of the dissimilarity coefficients are different, this order is unique [3]. There exist a class of so called Robinsonian matrices which admit always compatible orderings. Their elements do not decrease when moving away from the main diagonal in either direction. Robinsonian matrices appeared with Robinson (1951) for seriation problem in the field of archaeological science.

The methods of hierarchical cluster analysis transform the dissimilarity matrix D to an ultrametric matrix U . It is interesting to be noted that one ultrametric matrix admits 2^{n-1} compatible orderings (chains).

3 Principles of the Ordered Clustering algorithm

Let recall that the ascending hierarchical clustering procedures start with n distinct clusters of a single object which are the objects of X and the dissimilarity coefficients between all pairs of clusters are the elements of the initial intercluster dissimilarity matrix. At each level of the clustering process the two nearest clusters are joined together. Thus the number of clusters decreases with 1 and the dissimilarity between the new cluster and the other

existing clusters is calculated. The various hierarchical clustering methods imply different formulas of between cluster dissimilarity measure. The hierarchical process stops when are reached the desirable number of clusters or mainly when all objects are grouped in a single cluster. In the standard algorithm two clusters are united without taking into account the possibilities for orienting them in such a way that the closest object from the different clusters be adjacent.

In the ordering algorithms when at the level k of the hierarchical process C_1 and C_2 are the two nearest chains with their respective left and right ends o_{l1}, o_{r1} and o_{l2}, o_{r2} , there are four possibilities for linking these chains without changing the mutual arrangement of the objects in them. The order of the objects of C_1 and C_2 is fixed at lower levels. In the ordering algorithms of Brossier(1980) and Gruvaeus and Winer(1982) C_1 and C_2 are linked at their nearest ends. This ordering rule is applicable to all the hierarchical clustering methods. We use this principle in our algorithm for reordering the dendrograms obtained by the standard clustering procedures of the wide known statistical packages as BMDP and Statistica for Windows [14]. This algorithm uses as input only the $n(n - 1)/2$ dissimilarity coefficients of D and the amalgamation schedule of these procedures which is two dimensional array with $(n - 1)$ rows and three columns. The elements of the first two columns are numbers of two objects each one from the two clusters linked at the according level and the third column contain the respective amalgamation distances. The computer program in Pascal is furnished by use of a binary tree data structure. As a result of applying this reordering algorithm we obtain an ordered dendrogram which has the property that the neighbor terminal nodes (objects) and nonterminal nodes (clusters) in it are really close and vice versa.

Our original ordering algorithm called Ordering Single Link Algorithm (OSLA) uses a following definition for the dissimilarity between clusters.

Definition 3.1 *Let C_1 and C_2 be two chains with their respective left and right ends o_{l1}, o_{r1} and o_{l2}, o_{r2} . The dissimilarity d between C_1 and C_2 is*

$$d(C_1, C_2) = \min(\delta_{l1l2}, \delta_{l1r2}, \delta_{r1l2}, \delta_{r1r2}),$$

We introduce the following rule for connecting two chains:

If $\delta_{ij} = \min(\delta_{l1l2}, \delta_{l1r2}, \delta_{r1l2}, \delta_{r1r2})$, ($o_i \in \{o_{l1}, o_{r1}\}, o_j \in \{o_{l2}, o_{r2}\}$) the chains C_1 and C_2 are linked in such a way that their ends o_i and o_j to be adjacent in the new chain.

The advantages of the OSLCA can be summarized as follows.

- (i) It uses only the $n(n - 1)/2$ dissimilarity coefficients for the n objects. These dissimilarities are required and sorted once only and a hierarchical system of agglomerative clusters is easily constructed.
- (ii) It is particularly suitable for analysis of a dissimilarity matrix with missing elements. They are considered infinite.
- (iii) Its output is a unique dendrogram. The arrangement of the terminal nodes has some interesting mathematical and topological properties. This algorithm works fast and effectively.

There is a close relationship between OSLCA and the problem of the Minimum Spanning Tree (MST). The MST algorithm given by Gower and Ross (1969) aims to construct a connected graph which does not have any loops and contains all points(objects) and edges joining pairs of points. The length of this graph (called a spanning tree) is equal to the sum of the lengths of its edges and is to be minimal. When all edges of the graph are different the minimum spanning tree is unique.

Kruskal (1956) suggested that MST problem was related to the traveling salesman problem and Obruča (1968) took up this suggestion and described a technique of manipulating the MST solution in order to obtain a chain with length as small as possible which to be an approximate solution of the problem. An extension is considered with regard to the minimal wiring problem.

We will prove that under some conditions the obtained chain by the OSLCA will be a solution of the problem for constructing an unbranched spanning subtree (chain) with a minimal length.

4 Properties of OSLA

Let us consider an order of the observed n objects in a given chain O as a permutation of the numbers from 1 to n and denote by $A(O)$ the matrix of dissimilarity coefficients corresponding to this order. $A(O)$ is a symmetric matrix with non negative elements and the diagonal elements equal to 0. Note that the elements $a_{i,i+1}$ above the main diagonal are the dissimilarity coefficients between the neighbor objects in the chain and its length is equal to the sum of these elements $L(O) = \sum_{i=1}^{n-1} a_{i,i+1}$. The upper half of $A(O)$ is:

$$\begin{array}{cccccccc}
 a_{12} & a_{13} & \dots & a_{1i} & a_{1i+1} & \dots & a_{1n} & \\
 & a_{23} & \dots & a_{2i} & a_{2,i+1} & \dots & a_{2n} & \\
 & & & \dots & \dots & \dots & \dots & \\
 & & & & a_{i-1,i} & a_{i-1,i+1} & \dots & a_{i-1,n} \\
 & & & & & a_{i,i+1} & \dots & a_{in} \\
 & & & & & & \dots & \dots \\
 & & & & & & & a_{n-1,n}
 \end{array}$$

It is easy to note that in the rearranged dissimilarity matrix $A(O)$ according to the order obtained via OSLA the most similar pairs of objects tend to lie near the diagonal and the dissimilarity between the pairs tend to increase away from the diagonal. This organization helps to highlight the interobject relations and is very useful for examining large matrices. In the ideal case reordered matrix is Robinsonian but in practice such cases occur not very often. Here we consider several sets of inequalities for a dissimilarities of a chain. The goal is to find a "minimal" set of such inequalities which is sufficient for the chain to be of minimal length among all chains connecting the same objects.

Definition 4.1 *We will call a chain of n objects perfect if the dissimilarity coefficients between its members fulfill the following (call them overlapping) inequalities:*

$$a_{i,i+1} \leq a_{j,k}, (1 \leq j \leq i, i+1 \leq k \leq n). \quad (1)$$

Theorem 4.1 *The perfect chain has a minimal length.*

The proof of this theorem is given by Vandev et al.(1995b). The converse is generally false.

The perfect chains are useful because they are easy to check and have some natural nice properties.

Theorem 4.2 *Suppose there exist a perfect chain in the input dissimilarity matrix and all its elements are different, or at least all inequalities (1) are strict. Then the shortest chain is uniquely defined.*

Proof. The proof immediately follows the proof of theorem 4.1. Q.E.D.

Let us remind that in the OSLCA each cluster is a chain and at each level of the hierarchical process the two currently nearest chains are tied at their closest ends. Let us suppose now that $A(O)$ is the matrix corresponding to the order O of the objects obtained by the OSLCA. Each element $a_{i,i+1}$ is equal to one of the amalgamation levels $\{\Delta_1 < \Delta_2 < \dots < \Delta_{n-1}\}$ so that $\min\{a_{i,i+1}, 1 \leq i \leq n-1\}$ corresponds to the clustering level Δ_1 and so on and $\max\{a_{i,i+1}, 1 \leq i \leq n-1\}$ corresponds to the highest clustering level Δ_{n-1} .

Theorem 4.3 *Suppose we have two perfect chains to tie together. Suppose that*

- a. *the closest distance between both sets of objects is between objects which are closing elements in the both chains.*
- b. *this distance is not less than all distances between adjacent elements in the both chains.*

Then the resulting chain is perfect, i.e. has a minimal length.

Proof. The proof immediately follows by the definitions. Denote by o_1, o_2, \dots, o_p the elements of the first chain and by $o_{p+1}, o_{p+2}, \dots, o_n$ – the elements of the second one. So we have for any $1 \leq i \leq n-1$ that

$$a_{i,i+1} \leq a_{p,p+1} \leq a_{j,k}, 1 \leq j \leq p, p+1 \leq k \leq n.$$

On the other side if $1 \leq j, k \leq p$ or $p+1 \leq j, k \leq n$ then the same inequality follows from the perfectness of the starting chains for the corresponding values of i . Q.E.D.

Note that in an hierarchical nearest neighbor procedure the last condition of the theorem 4.3 need not to be checked – this condition is fulfilled automatically. More over the following theorem may be stated:

Theorem 4.4 *Suppose there exist a unique perfect chain in the input data. Then the hierarchical nearest neighbor procedure in the form of OSLCA will find it.*

Proof. The proof immediately follows by the fact that at given stage of the procedure all constructed chains are subchains of the existing perfect chain. Then one has to check the distance $a_{j,k}$ which comes next in the sorted array of distances. It turns out that if j and k belong to different chains then neither of them may be "interior" because some of the ends of the corresponding chains are to be closer. Q.E.D.

It is interesting the question of compatibility between dissimilarity matrix D and the chain of terminal nodes obtained by OSLA. It is obvious that if D is a Robinsonian matrix, i.e. admit compatible order, so OSLA will find this order and the correspondent chain is perfect. The proof is too elementary if recall the properties of the Robinsonian matrices. Chepoi and Fichet (1996) present an algorithm for testing if one dissimilarity matrix admit compatible order. The opposite assertion, is not true in general. Thus the class of the dissimilarity matrices inducing a perfect order are wider than the class of Robinsonian matrices. Here is an example of a perfect matrix which is not Robinsonian, i.e. doesn't admit a compatible order.

$$\begin{array}{cccc} 1 & 7 & 5 & 9 \\ & 4 & 8 & 6 \\ & & 3 & 10 \\ & & & 2 \end{array}$$

5 Examples

Here we demonstrate two examples and compare usual dendrograms with ordered ones.

Example 1. The experiment was conducted in two adjacent years with young bulls from four breeds during the age of 15 to 18 months, divided in nine groups. The breeds are : Poll Hereford (PH), Native Gray cattle (NG), Black and White cattle (BW) and cross between Poll Hereford and Native Gray cattle (F1). The various groups of cattle were fattened in two ways - stall (O), stall and grassing (K). The names of the groups are PH-O1, NG-O1, BW-O1, BW-K1 of the first year and PH-O2, PH-K2, NG-O2, NG-K2, F1-K2. Thus the name PH-O1 means the cattle group of Poll Hereford, stall fattened in the first year.

Six attributes characterizing the meat yield were measured - dressing percentage, carcass percentage of bones, carcass percentage of total meat amount, carcass percentage of the various meat quality. The means of these attributes were obtained for each group.

The nine groups were considered as objects for cluster-analyzing. As a dissimilarity measure between objects is used the Euclidean distance. Table 1 contains the upper half of the dissimilarity matrix according to the ordering obtained by OSLA:

A dissimilarity matrix for 9 objects								
Breeds	NG-O1	BW-K1	BW-O1	F1-K2	NG-K2	NG-O2	PH-K2	PH-O2
PH-O1	3.93	4.83	6.45	9.78	9.47	7.82	7.06	7.30
NG-O1		3.68	5.97	7.33	6.86	5.86	8.23	9.28
BW-K1			2.89	6.32	6.37	5.48	6.49	6.63
BW-O1				5.10	5.67	5.80	6.89	6.50
F1-K2					1.49	3.56	7.24	7.68
NG-K2						2.61	6.70	7.87
NG-O2							4.83	6.45
PH-K2								2.49

Table 1

The Single Link (SL) clustering method is applied to the distance matrix using the standard SL procedure of the package STATISTICA for Windows and the arranging OSL algorithm.

Using the hierarchical cluster-analysis we aim to ascertain whether the similarity between groups on the base of the meat performance traits is related to the breed or the way of fattening, or to the forage quality (year).

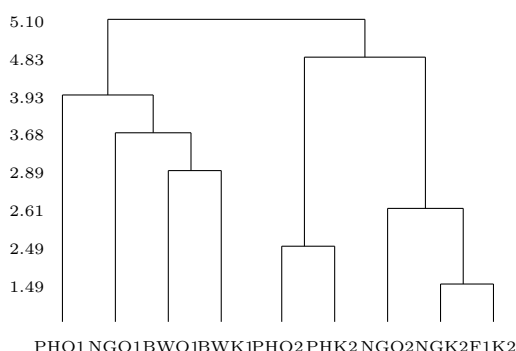


Fig.2

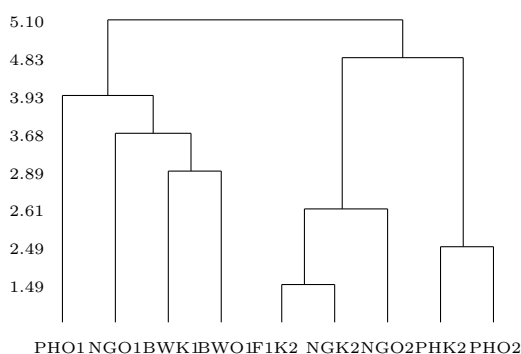


Fig.3

The dendrograms obtained by the two SL procedures are shown in Fig.2 and Fig.3. Two main clusters are formed into both hierarchical schemes. The first one comprises the cattle groups observed in the first year of the experiment and the second cluster consist of the groups of the second yea. The conclusion from this clustering is that the cattle fattened during the same year what means the same quality of the forages have similar meat performance.

Other clusters consisting of comparatively close cattle groups are (PH-K2, PH-O2) and (BW-K1,BW-O1). The conclusion is that the cattles from the same breed fattened through the same year have similar meat performance.

Very interesting cluster is (NG-O2,NG-K2,F1-K2). The group F1-K2 of cattles crosses between native Gray cattle and Poll Hereford show stronger influence of the mother breed (NG) on the investigated traits which is unexpected fact. The reason for this is probably the way of fattening - stall and grassing.

Let us compare the two dendrograms above - the first obtained by the standard SL algorithm without taking into account the possibility of ordering the objects and the second one is the result of the rearranging algorithm.

Although the clustering structure in both dendrograms is identical it is easy to see from Table 1 that the order of the objects in Fig.3 gives better presentation of the relationships

between the investigated cattle groups. For instance the group NG-O1 is closer to BW-K1 than to BW-O1. Then the group BW-K1 is more similar to F1-K2 than to PH-O2. It is clear that the ordered SL algorithm gives more accurate visual representation of the dissimilarities between the observed objects.

Anoder interesting result is that the resulting via OSLA chain is perfect, i.e. it is with minimal length (see Table 1).

Example 2. In the second example are used data from 305d lactation of Black and White cows from three herds from the same region over the period from 1981 to 1995. The initial data contained records of the following attributes: the number of lactation (1,2,3); milk yield during the correspondent lactation (in kg); average percent of the milk fat, durability of days open (in days); age of the first calving (in days). The individuals are separated in groups according to the number of lactation and the level of milk yield. There were nine milk yield levels formed on the base of the milk yield- the first level comprises the individuals with milk yield to 3000kg, the adjacent levels are with increase interval of 1000kg (great than 3000 and less or equal to 4000kg and so on). In this way are obtained 20 groups. For each of them are calculated the group means of the milk fat percentage, days open and age of the first calving.

The distance measure between cases (20 groups) is the Euclidean distance using standardized data. The Unweighted Pair-Group Method using arithmetic averages (UPGMA) was applied to the distance matrix as one of the most widely used clustering methods.

The aim of the investigation is to establish to what degree the similarity between the groups based on the reproductive performance, milk concentration and the growth of the cows as heifers is related to the milk level or to the number of the lactation.

The results from the standard UPGMA and our reordering algorithm are visualized by the dendrograms in the corresponding Fig. 4 and 5. Although the vertical structure of the both dendrograms is merely identical the difference between them is in the arrangement of the terminal nodes in them. In these dendrograms the observed groups are denoted as the first number means the number of lactation and the second number - the milk level.

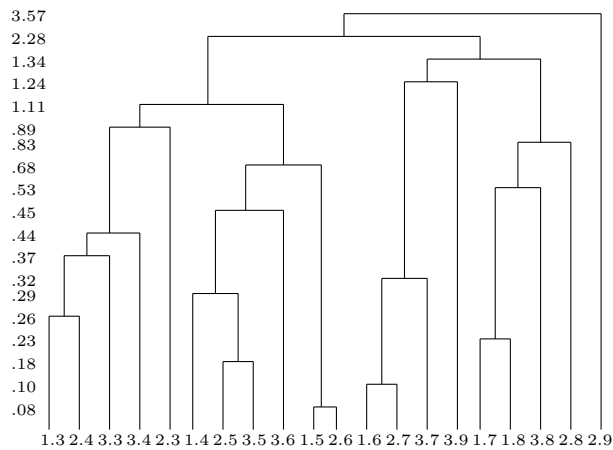


Fig.4

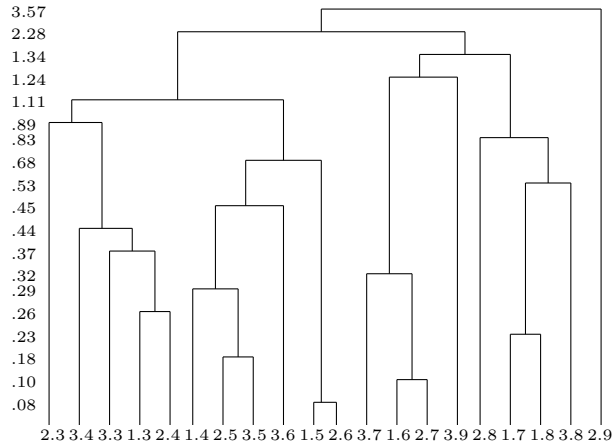


Fig.5

The hierarchical clustering structure of the both dendrograms present two main clusters, one of them consisting of groups with milk level from to 6000kg and various lactations and the second one comprise groups with milk yield from 7000 to 9000 kg and various lactations. Into each one of these main clusters are formed two subclusters in which are united groups with milk yield to 4000kg, from 5000kg to 6000kg, from 6000kg to 7000kg and from 8000kg to 9000kg. The conclusion is that the similarity on the base of milk fat percentage, days open and age of the first calving is due rather to the milk level than to the number of lactation. The interesting exception to this trend are the groups of the first lactation. Each one of these groups 1.3, 1.4, 1.5, 1.6, 1.7, 1.8 is included in clusters comprising groups of second and third lactation with the same or higher milk level (e.g. 1.3 with 2.4; 1.4 with 2.5 and 3.5). This trend is better presented in by the rearranged dendrogram and has reasonable explanation - the young cows are close to the cows with higher milk yield in a sense of the milk fat percentage and the reproductivity. The isolated cluster 2.9 consist of cows of the second lactation with over 8000kg milk yield.

For this research application we found relation between the obtained classification and the milk productivity of the cows.

6 Discussion

It is clear that the OSLA produces a dendrogram which gives more accurate visual representation of the hierarchical clustering structure and of the relationships between the observed objects than the usual single link analysis.

Although the OSLCA yields undoubtedly good results the arrangement of the objects in the final chain is not always perfect.

Also our algorithm does not give always an optimal solution of the minimal length problem because of its rules for constructing of the chains (without looping and connecting only in the ends). It is apparent that according to these restrictions some of the range ordered dissimilarity coefficients will be skipped as their corresponding objects are internal elements of two distinct chains or belong to the same chain. The number of such omitted dissimilarities depends on the concrete data.

Let n_b denote the number of all skipped dissimilarities between pairs of objects belonging to different chains where at least one of them is an internal element of a chain. Let n_w be the number of all skipped dissimilarities between pairs of objects belonging to the same chain and n_s be the number of all scanned sorted dissimilarities from the beginning to the end of the hierarchical process. It is clear that $n_s = n_b + n_w + n - 1$ where $n - 1$ is the number of used dissimilarities. The value of n_w does not influence on the perfectness of the chains. As an indicator of unperfectness (UP) of the chain obtained by OSLCA, we introduce the following proportion $UP = \frac{n_b}{n_b + n - 1}$. For the perfect chain UP is 0.

There are many ways of defining a criterion which measures if an order on X is more or less compatible with the dissimilarity matrix D .

Let O is a chain containing N elements, T is the set of all ordered triples of O and S is a set of all elements of T compatible with D (see Definition 2.3). We propose the following criterion:

$$C(O, D) = \frac{|S|}{|T|}, \text{ where } |T| = \binom{N}{3}.$$

For the ordering in Example 1 $C(O, D) = .75$.

Acknowledgment

Authors are grateful to Dr Z. Gergovska and Dr R. Otusbirov from the Dept. of Cattle Husbandry and to R. Otusbirov from Dept. of Meat and Meat Products of the Thracian University for permission to use their data, as well as for helpful discussions on the practical interpretation of the obtained results.

References

- [1] Benzecri J. P. (1965), Problemes et methodes de la taxonomie, Raport de recherche de l'Universite de Rennes, France.
- [2] Benzecri J. P. (1973), L'analyse des donnees. 2, La taxonomie, Dunod, Paris.
- [3] Brossier G. (1980), Représentation ordonnées des classifications hiérarchiques, *Statistique et Analyse des Données*, 2, pp. 31-44.
- [4] Brossier G. (1982), Classification á partir de matrices carrées non symetriques, *Statistique et Analyse des Données*, v.7, No 2, pp.22-40.
- [5] Brossier G. (1983), *Ordonnement des hiérarchies*. 3^e journées internationales. Analyse des données et informatique. INRIA, Versaille.
- [6] Brossier G. (1984), Algorithmes d'ordonnement des hierarchies binaires et proxetes, *Revue de Statistique Appliquee*, vol. 32, No 3, rp. 65-79.
- [7] Chepoi V., Fichet B. (1996), Recognition of Robinsonian dissimilarities, to appear

- [8] Degerman R. (1982), Ordered bynary trees constructed through an application of Kendall's tau, *Psychometrika*, vol.47, No 4, pp. 523-527.
- [9] Gower, G.S. and Ross, G.J.S. (1969), Minimum spanning trees and single-linkage cluster analysis, *Applied Statistics*, **18**, 54–64.
- [10] Gruvaeus G., Wainer H. (1972), Two additions to hierarchical cluster analysis, *British Journal of Mathematical and Statistical Psychology*, 25, pp. 200-206.
- [11] Kruskal, J.B. (1956), On the Shortest spanning subtree of a graph and the travelling salesman problem, *Proceedings of the American Mathematical Society*, **7**, 48 – 50.
- [12] Obruča, A.K. (1968), Spanning tree manipulation and the travelling salesman problem, *Computer Journal*, **10**, 374–377.
- [13] STATISTICA for Windows, Vol.1, 3, General conventions, StatSoft, Inc., 1994
- [14] Vandev D. L. and Tsvetanova Y.G. (1995a), Ordered Dendrogram, paper presented at the 9th European Meeting of the Psychometric Society, Leiden, the Netherlands, July 4-7.
- [15] Vandev D. L. and Tsvetanova, Y.G. (1995b) , Perfect Chain and Single Linkage Clustering, *Proceedings of SDA '95*, Varna.