Another, nonparametric measure of association is Goodman and Kruskal's γ, defined as $(S_+ - S_-)/(S_+ + S_-)$, where S_+ and S_- are the number of concordances and discordances, respectively. A concordance or discordance in the context of matrix comparison is defined by comparing each pair of pairs. For example, the pairs h_{12} and h_{14} in **H** and d_{12} and d_{14} in $\mathbf{D}_1$ are discordant because $2 < 5$ in **H** and $2 < 10$ in D_1. For these data, γ is 1.0.

Further information on dendrogram comparison and some applications are given in Chapter 9.

4.4.3 Mathematical properties of hierarchical methods

A number of mathematical properties can be defined for clustering methods. One of these, the *ultrametric property,* was first introduced by Hartigan (1967), Jardine *et al.* (1967) and Johnson (1967), and has since been shown to be related to various features of clustering techniques, in particular the ability to represent the hierarchy by a dendrogram. The *ultrametric property* states that

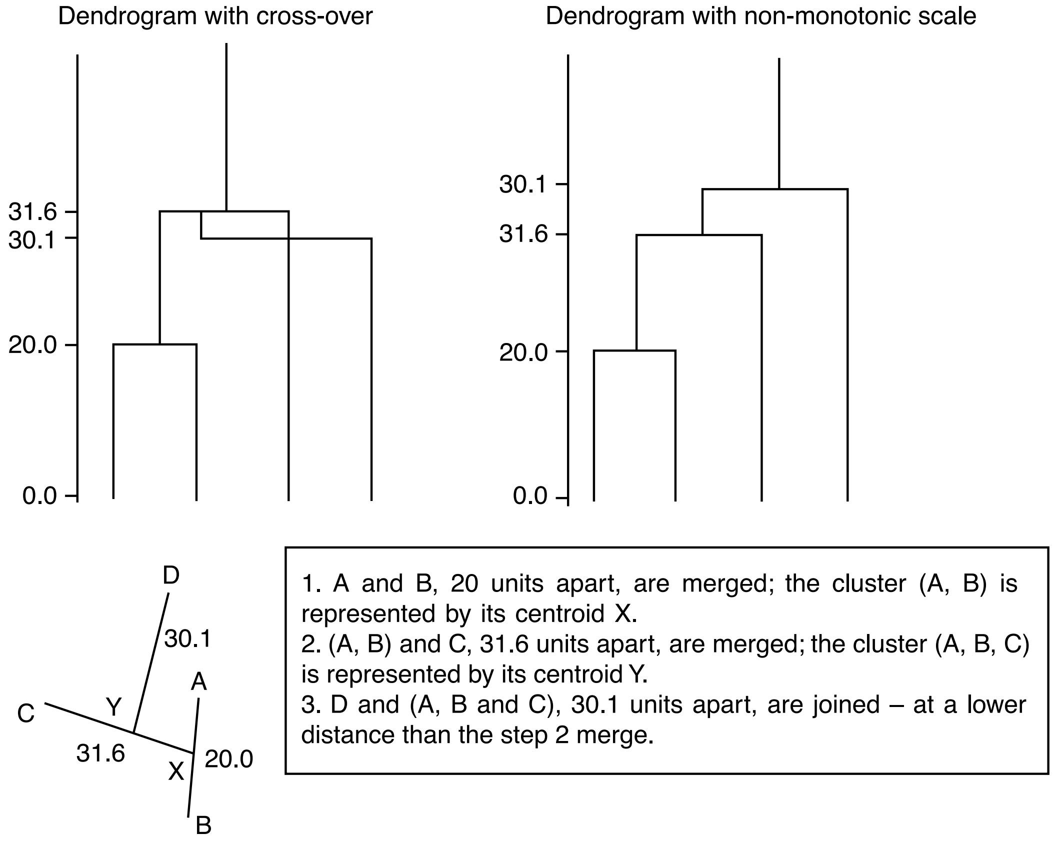
$$h_{ij} \leq \max\left(h_{ij}, h_{jk}\right) \text{ for all } i, j \text{ and } k, \tag{4.12}$$

where h_{ij} is the distance between clusters i and j. An alternative way of describing this property is that for any three objects, the two largest distances between objects are equal. The property does not necessarily (or even usually) hold for the elements of proximity matrices. However, it does hold for the heights h_{ij} at which two objects become members of the same cluster in many hierarchical clustering techniques.

A consequence of failing to obey the ultrametric property is that *inversions* or *reversals* can occur in the dendrogram. This happens when the fusion levels do not form a monotonic sequence, so that a later fusion takes place at a lower level of dissimilarity than an earlier one. Morgan and Ray (1995) describe some empirical studies of inversions for a number of methods. Inversions are not necessarily a problem if the interest is in one particular partition rather than the complete hierarchical structure. They may also be useful in indicating areas where there is no clear structure (Gower, 1990). However, as Murtagh (1985) points out, reversals can make interpretation of the hierarchy very difficult, both in theoretical studies of cluster properties and also in applications where a hierarchical structure is an intrinsic part of the model. This is because the nested structure is not maintained, as shown in Figure 4.8. Both centroid and median clustering can produce reversals.

A related feature of clustering methods is their tendency to 'distort' space. The 'chaining' effect of single linkage, in which dissimilar objects are drawn into the same cluster, is an example of such distortion, *space contraction* in this case. The opposite type of distortion, where the process of fusing clusters tends to draw clusters together, is *space dilation,* as found in complete linkage. *Space-conserving* methods, such as group average linkage, obey the following inequality:

$$d_{iuv} \leq d_{i(uv)} \leq D_{iuv}, \tag{4.13}$$


Dendrogram with cross-over
31.6
30.1
20.0
0.0
Dendrogram with non-monotonic scale
30.1
31.6
20.0
0.0
D
30.1
A
C
Y
31.6
X
20.0
B

1. A and B, 20 units apart, are merged; the cluster (A, B) is represented by its centroid X.
2. (A, B) and C, 31.6 units apart, are merged; the cluster (A, B, C) is represented by its centroid Y.
3. D and (A, B and C), 30.1 units apart, are joined – at a lower distance than the step 2 merge.

Figure 4.8 *Example of an inversion or reversal in a dendrogram.* (Adapted with permission of the publisher, Blackwell, from Morgan and Ray, 1995.)

where d_{iuv} and D_{iuv} are the minimum and maximum distances between object i and clusters u and v, respectively, and $d_{i(uv)}$ is the distance between object i and the fusion of clusters u and v. In other words, distances to merged clusters are intermediate between distances to the constituent clusters. Space-conserving methods can be thought of as 'averaging' the distances to clusters merged, while space-dilating (-contracting) methods move the merged clusters further from (closer to) each other.

A number of admissibility properties were introduced by Fisher and Van Ness (1971). Such properties would be desirable qualities, other things being equal, and as such they can aid in the choice of an appropriate clustering method. One of these, (*k-group*) *well-structured admissibility,* has been related to the Lance and Williams parameters by Mirkin (1996), who terms it *clump admissibility.* (There are a number of other subtypes of well-structured admissibility, but this one relates directly to space conservation and the ultrametric condition.) Mirkin defines this property as follows:

- *Clump admissibility*: there exists a clustering such that all within-cluster distances are smaller than all between-cluster distances.