***dm* = Matching**

This metric is the same as Hamming, if Hamming is given only boolean data, that is, only ones and zeroes.

$$d(u,v) = \frac{b+c}{D}$$

HCluster Linkage Calculation Methods

You use the HCluster /LINK=*linkMethod* flag to specify the method used to determine the dissimilarity between nodes in the dendrogram that represent more than one data vector. This is also referred to as the "linkage" method. Our definitions of node dissimilarity follows Python scipy.cluster.hierarchy.linkage.

linkMethod is a keyword identifying the method to use. Each of these keywords, described below, selects a method for measuring the dissimilarity between two clusters, *A* and *B*, with individual dissimilarities *a* and *b*. The symbol | A | denotes the number of elements in a cluster. It is not always possible to give an expression in this form.

Alternately, dissimilarity between a new node and other nodes in the tree can be computed from the dissimilarities of the two nodes being combined. That is, if nodes I and J are combined to form K, then it is possible to compute the dissimilarity from K to any other node L in terms of I and J.

Each of the following linkage method descriptions includes two equations. The first describes the method for computing dissimilarity between nodes given the dissimilarities of each of the component dissimilarities. The second describes the method for computing the dissimilarity to another node from a node that combines I and J.

The following values are supported for *linkMethod* keyword. If you omit /LINK, HCluster defaults to the average method.

***linkMethod* = single**

This is the default dissimilarity metric.

$$d(A,B) = \min_{a \in A, b \in B} d(a,b)$$

That is, the dissimilarity is the dissimilarity between the two nearest elements of each cluster. Also called Nearest Point algorithm.

$$d(K,L) = min(d(I,L), d(J,L))$$

***linkMethod* = complete**

$$d(A,B) = \max_{a \in A, b \in B} d(a,b)$$

Also called Farthest Point algorithm.

$$d(K,L) = max(d(I,L), d(J,L))$$

***linkMethod* = average**

$$d(A,B) = \frac{1}{|A|\,|B|} \sum_{a \in A, b \in B} d(a,b)$$

That is, the average of all the pairwise dissimilarities between points in each cluster.