Comparision Results

Parismita Das

5 March 2017

Comparision between rioja and adjClustBand_heap

From the Merge Data Table and Dendrogram of rioja and adjClustBand_heap, We compare the clustering values.

Rioja Data Table considering distance matrix (dissimilarity matrix) as the same given in the file dissimilarity2.txt.

	V1	V2	height
1	-30	-31	0.01
2	-81	-82	0.02
3	-28	-29	0.03
4	-96	-97	0.04
5	-3	-4	0.07
6	-127	-128	0.10

From the table we can say that the (30,31) merges at height 0.01, (81,82) at 0.02 and so on.

If we consider the dissimilarity matrix such that it is 2-2*s', where s' is scaled similarity matrix. We get

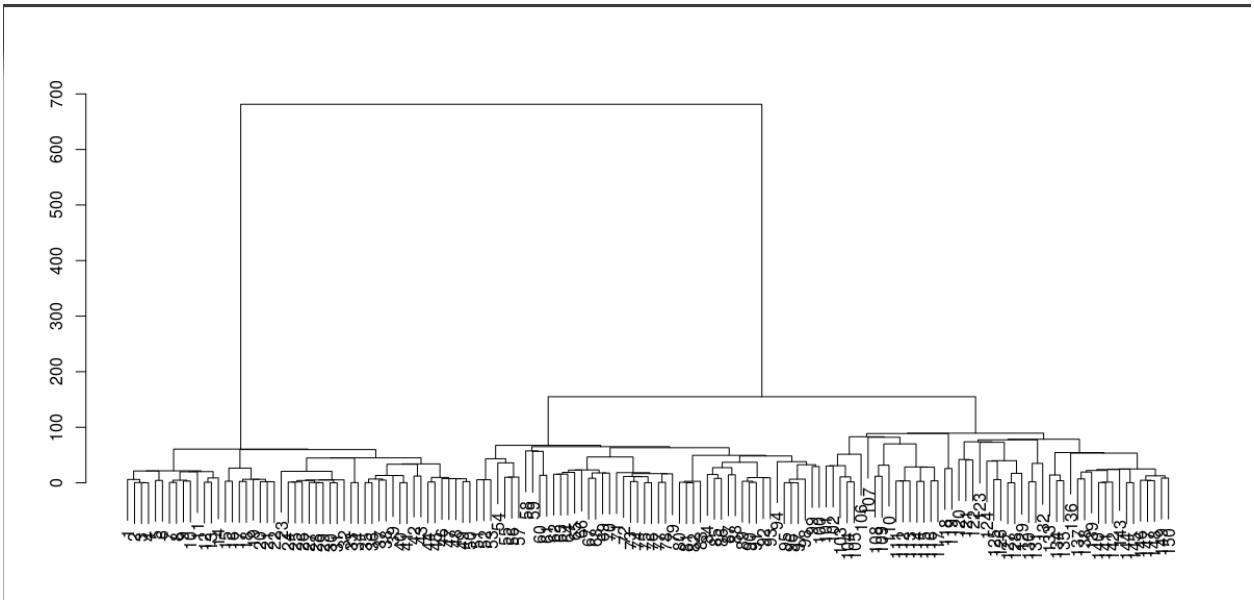
	V1	V2	height
1	-28	-29	0.001064484
2	-30	-31	0.002372169
3	-40	-41	0.004869267
4	-3	-4	0.007599009
5	-35	-36	0.011270532
6	-23	-24	0.015951043

	V1	V2	gain	height
1	-28	-29	0.001064484	0.01
2	-30	-31	0.001307686	0.02
3	-40	-41	0.002497098	0.03
4	-3	-4	0.002729742	0.04
5	-35	-36	0.003671523	0.07
6	-23	-24	0.004680510	0.10

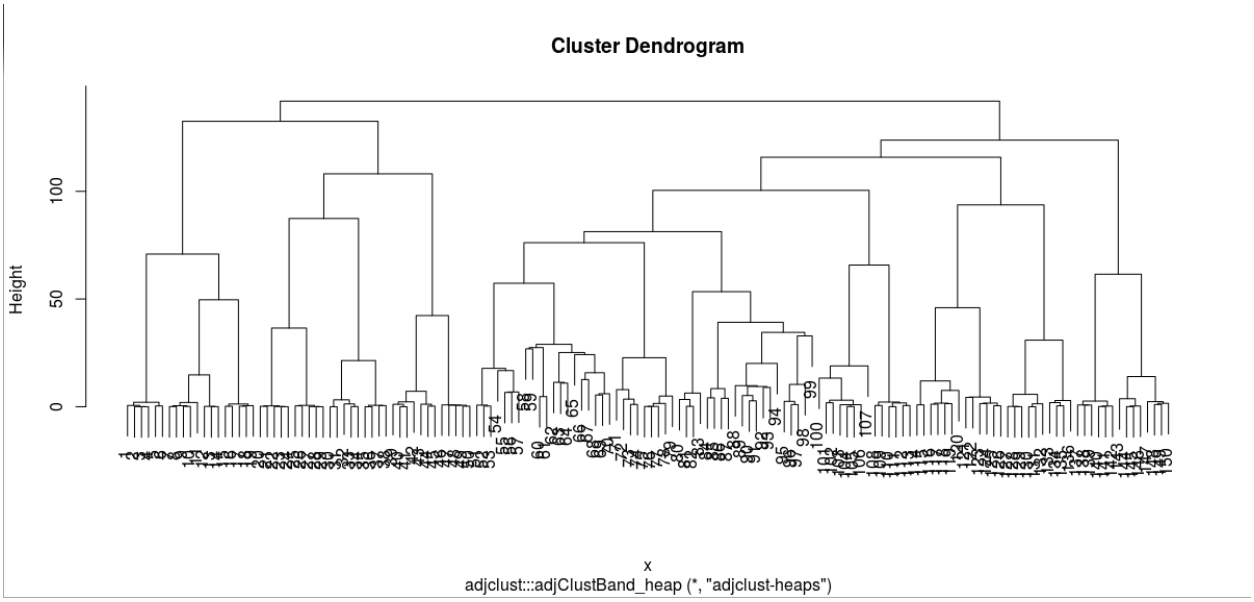
adjClustBand_heap Data Table with maximum band similarity as shown above.

From the table we can say that the merge data of both the functions are same when we take dissimilarity matrix as 2-2*s' and maximum band of similarity matrix for adjBandClust_heap

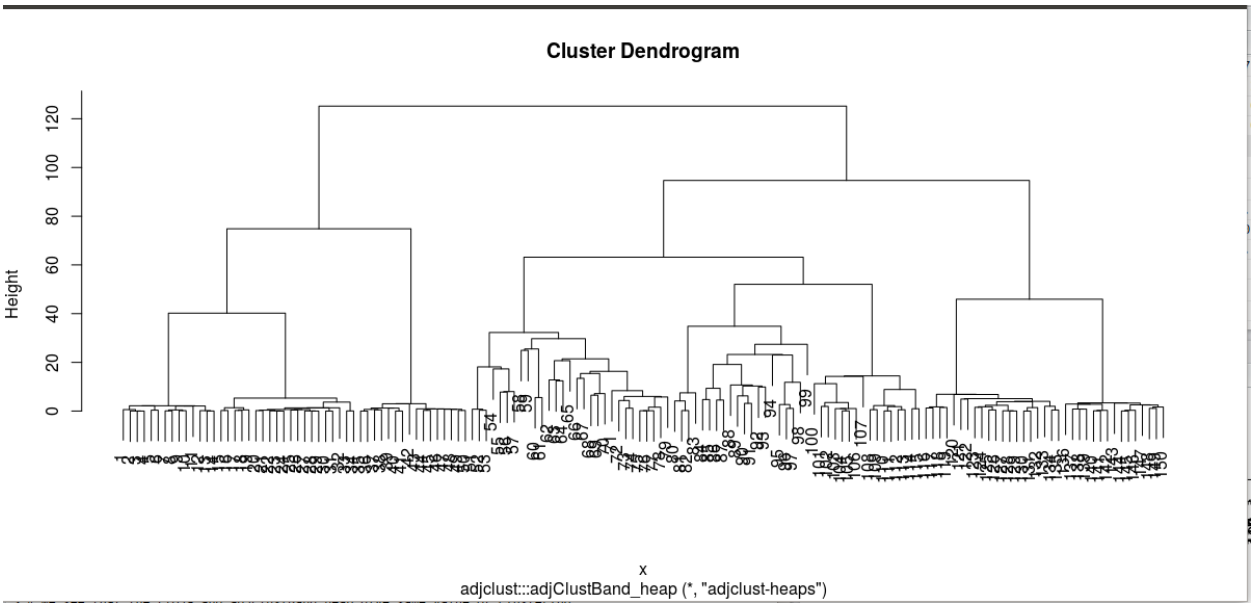
The Dendrogram obtained from rioja package, as we consider the dissimilarity matrix as the same given in the file dissimilarity2.txt.



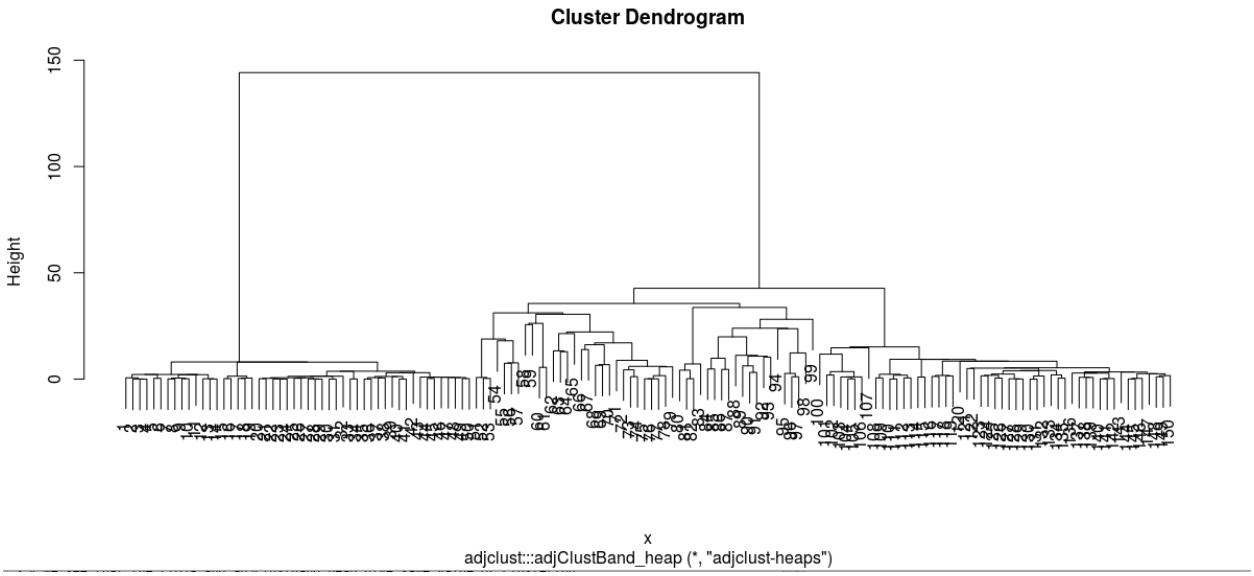
Rioja Dendrogram



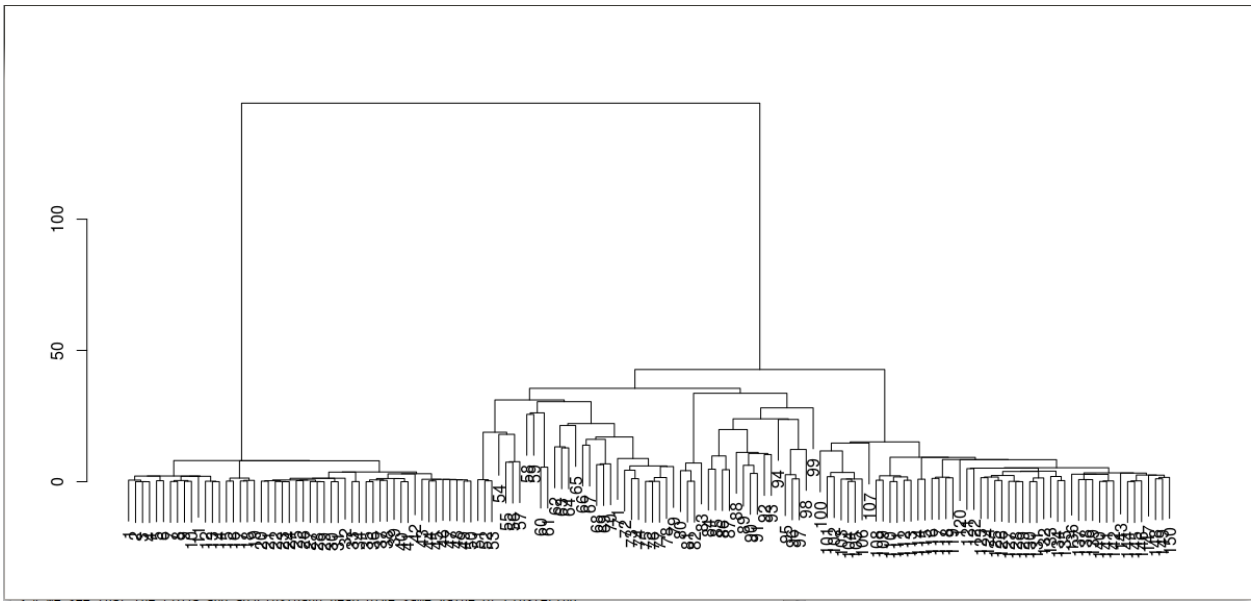
adjClustBand_heap Dendrogram for h = 5



adjClustBand_heap Dendrogram for h = 20



adjClustBand_heap Dendrogram for h = maximum height of the matrix



Rioja Dendrogram for dissimilarity as 2-2*s'

By manipulating the similarity matrix such that s" = 2-2*s' we see that the rioja gives the same plot as adjClustBand_heap when full band is considered.

After normalising the matrix such that the diagonal elements are 1, the Cluster Merging details that are comming are almost similar as of rioja cluster merge data. Hence we can conclude that the adjClustBand_heap gives same result as that of rioja thus proving its correctness and uses a more optimised algorithm.