DELTA classification programs

9 February 2019

M.J. Dallwitz

Contents

Introduction

The following programs work on distance matrices produced by the program Dist.

Pclass. Agglomerative clustering by combinatorial sorting strategies.
Nsim. Lists nearest neighbours of taxa.

To run them, open a 'Command Prompt' window (see Installing and running the programs of the DELTA System), and change to the relevant data folder (using the standard Windows command 'cd' or the DELTA command 'to'). Enter the program name and respond to the prompts, which request information such as the number of OTUs (taxa) and the name of the distance-matrix file.

Pclass

Pclass carries out agglomerative clustering by combinatorial sorting strategies.

The prompts, and an example of user input for the DELTA sample data, are shown below.

```
Enter number of OTUs:14
Enter name of file with distances:grass.dis
Print output on screen, file, or both [default F]:
Enter name for print file [default PCLASS.OUT]:
Trees in Txplot, Nexus, Phylip, or Hennig format ("-" for none) [default T]:p
Enter name for tree file [default grass.TRE]:
Enter name of taxon-name file, or "-" if none [default grass.NAM]:

Available methods are:
   1 - nearest neighbour   2 - furthest neighbour
   3 - UPGMA               4 - WPGMA
   5 - L & W flexible      6 - Colless variable
   7 - ISS flexible        8 - ISS variable
   9 - generalized
Enter method number [default 7]:
Enter clustering intensity [default 0.25]:
Rerun? [default N]:
```

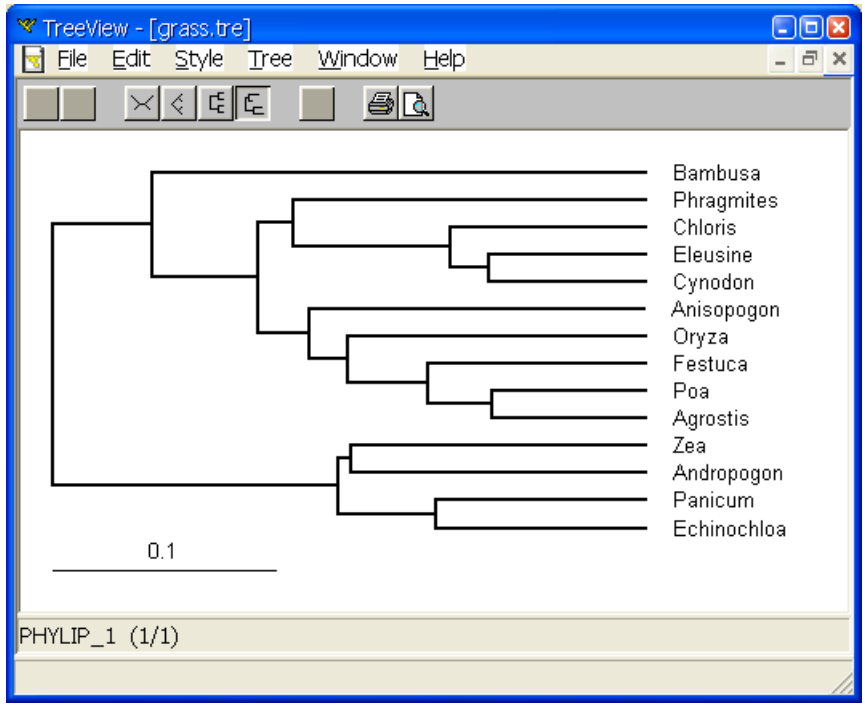
Most of the clustering methods are described in 'Sneath, P.H.A., and Sokal, R.R. 1973. Numerical taxonomy — the principles and practice of numerical classification. W. H. Freeman: San Francisco'. The 'ISS flexible' method is described in 'Dallwitz, M.J. 1988. A flexible clustering method based on UPGMA and ISS'.

Output produced by Pclass can be converted to graphical trees by means of the program TreeView (web.archive.org/web/20180714061902/http://taxonomy.zoology.gla.ac.uk/rod/treeview.html (14 July 2018)). When running Pclass, specify 'Phylip' format for the tree.

The output (pclass.out) from the above run is shown below.

```
PCLASS Version 2.02 (MS-DOS). 28-AUG-16

Analysis of data from file grass.dis

Run number 1.

Increment-in-sum-of-squares flexible sorting, intensity 0.25

     Fusion                      Level
   1  +   13  =   15    at     0.1366
   6  +    8  =   16    at     0.1401
   5  +   16  =   17    at     0.1743
   7  +   11  =   18    at     0.1869
   9  +   15  =   19    at     0.1941
   2  +   14  =   20    at     0.2633
  10  +   19  =   21    at     0.2666
  18  +   20  =   22    at     0.2745
   3  +   21  =   23    at     0.3005
  12  +   17  =   24    at     0.3158
  23  +   24  =   25    at     0.3469
   4  +   25  =   26    at     0.4418
  22  +   26  =   27    at     0.5317
```

A TreeView display of the above data is shown below. Note that annotation for the length of the scale line should be doubled to measure the fusion level; in this example, the scale annotation should be 0.2.



Nsim

Nsim lists nearest neighbours of taxa.

The prompts, and an example of user input for the DELTA sample data, are shown below.

```
Enter number of OTUs:14
Enter name of file with distances:grass.dis
Is there a file of OTU names? [default Y]:
Enter name of file with OTU names [default grass.NAM]:
Output on screen, file, or both [default F]:
Enter line length for output [default 80]:
Enter name for output file [default NSIM.OUT]:
Enter number of neighbours required [default 13]:6
Enter name or number of reference OTU (default all):
```

The output (nsim.out) from the above run is shown below.

```
NSIM Version 2.03 (MS-DOS). 28-AUG-16

Table of 6 nearest neighbours in matrix grass.dis

From:          To:

Agrostis....   Poa......... 0.1366  Festuca..... 0.2123  Anisopogon.. 0.2364
               Cynodon..... 0.2625  Panicum..... 0.2678  Chloris..... 0.2798

Andropogon..   Echinochloa. 0.2358  Zea......... 0.2633  Panicum..... 0.2929
               Agrostis.... 0.3414  Chloris..... 0.3521  Cynodon..... 0.3881

Anisopogon..   Agrostis.... 0.2364  Festuca..... 0.2509  Poa......... 0.2899
               Chloris..... 0.3438  Oryza....... 0.3641  Bambusa..... 0.3868

Bambusa.....   Oryza....... 0.3030  Festuca..... 0.3389  Phragmites.. 0.3695
               Anisopogon.. 0.3868  Poa......... 0.3887  Cynodon..... 0.4458

Chloris.....   Cynodon..... 0.1596  Eleusine.... 0.1821  Poa......... 0.2107
               Festuca..... 0.2668  Phragmites.. 0.2688  Agrostis.... 0.2798

Cynodon.....   Eleusine.... 0.1401  Chloris..... 0.1596  Poa......... 0.2145
               Echinochloa. 0.2439  Agrostis.... 0.2625  Panicum..... 0.2728

Echinochloa.   Panicum..... 0.1869  Andropogon.. 0.2358  Cynodon..... 0.2439
               Zea......... 0.2439  Chloris..... 0.2805  Phragmites.. 0.3329

Eleusine....   Cynodon..... 0.1401  Chloris..... 0.1821  Poa......... 0.1843
               Festuca..... 0.2667  Agrostis.... 0.2834  Phragmites.. 0.3174

Festuca.....   Poa......... 0.1645  Agrostis.... 0.2123  Oryza....... 0.2371
               Anisopogon.. 0.2509  Phragmites.. 0.2607  Eleusine.... 0.2667

Oryza.......   Festuca..... 0.2371  Poa......... 0.2375  Agrostis.... 0.2849
               Cynodon..... 0.2858  Phragmites.. 0.2901  Bambusa..... 0.3030

Panicum.....   Echinochloa. 0.1869  Agrostis.... 0.2678  Cynodon..... 0.2728
               Zea......... 0.2837  Andropogon.. 0.2929  Phragmites.. 0.2933

Phragmites..   Festuca..... 0.2607  Poa......... 0.2633  Chloris..... 0.2688
               Oryza....... 0.2901  Panicum..... 0.2933  Cynodon..... 0.2978

Poa.........   Agrostis.... 0.1366  Festuca..... 0.1645  Eleusine.... 0.1843
               Chloris..... 0.2107  Cynodon..... 0.2145  Oryza....... 0.2375

Zea.........   Echinochloa. 0.2439  Andropogon.. 0.2633  Panicum..... 0.2837
               Agrostis.... 0.3520  Oryza....... 0.3811  Cynodon..... 0.3940
```