

Welcome to the WEITZPRO Package

Calculation of the Weitzman diversity function

Written by Serguei Derban (HU Berlin/ ENSAE Paris)
in collaboration with Jean-Louis Foulley and Louis Ollivier
(INRA-SGQA Jouy-en-Josas), 2002.

REFERENCE

Derban S., Foulley J.-L., Ollivier L., 2002. WEITZPRO: a software for analysing genetic diversity. INRA, Paris.

COPYRIGHT NOTICE

(c) Copyright 2002 by Serguei Derban, Jean-Louis Foulley, Louis Ollivier and the INRA. Permission is granted to copy this document provided that no fee is charged for it and that this copyright notice and the disclaimer below are not removed.

INTERDEPOSIT CERTIFICATION: IDDN.FR.001.040038.000.R.P.2003.000.20700

DISCLAIMER

The authors and the INRA can not be held responsible for the correctness or quality of the programmes provided. Liability claims regarding damage caused by the use of any components of this package, including loss of information, will therefore be rejected.

REFERENCES

Weitzman M. L., On diversity, Quart. J. Econ. 107 (1992) 363-405.

Weitzman M. L., What to preserve? An application of diversity theory to crane conservation, Quarter. J. Econ. 108 (1993) 157-183.

Thaon d'Arnoldi C., Foulley J.L., Ollivier L., An overview of the Weitzman approach to diversity, Genet. Sel. Evol. 30 (1998) 149-161.

CONTENTS OF THIS PACKAGE

1 README.txt -- this description of the contents of the package

The compiled programs for Windows

2 weitzman.exe -- computes the diversity using the direct algorithm, makes the table of representatives and links and shows the location of the leaves in the classification tree

3 wmarg.exe -- computes the marginal loss of diversity for each member of set, i. e. its contribution to diversity

4 wgroup.exe -- computes the marginal loss of diversity for each predefined group, i. e. its contribution to diversity.

USING THE PROGRAMMES

The programmes were compiled as Win32 console application and need no special installation. You just need to extract them from the downloaded zip-file into a directory, for example C:\WEITZMAN. The programmes can be run from the Command-line Windows (MS-DOS Windows) or by double-clicking on the program icon.

Before running the programs from Command-line Windows you should move into the directory where the programmes were copied in by executing the following command:

```
cd c:\WEITZMAN ("WEITZMAN" is the name of directory you have chosen)
```

All the programmes take some arguments. If you use a programme without any argument, the program will invite you to enter them. The first argument is common for all the programmes. It is the name of text file containing a square distance matrix (input file). This file must have the following form: the dimension of the distance matrix in the first line, the distance matrix in the following lines, including the names of the elements of the matrix in the first column. The elements in lines should be separated by tabs or spaces. If the lines are too long, you can break them. The names of elements must not exceed 10 symbols.

Example: matrix.txt

```
-----
6
DEBB01    0    2    8    4    7    1
GBLR03    2    0   14    3   1.5  3.5
NOLR05    8   14    0   11   10    6
ISLR09    4    3   11    0   0.5   12
PLPU01    7   1.5  10   0.5    0    5
CZPR01    1   3.5   6   12    5    0
-----
```

Attention: this must be a simple text file! (No Word files!)

You should enter the name of input file with its extension (matrix.txt and not just matrix)!

DESCRIPTION OF THE PROGRAMMES

weitzman.exe

The programme "weitzman.exe" computes the Weitzman diversity function $V(S)$ using the direct algorithm and makes the table of representatives and links. It takes only one argument, the name of text file containing a square distance matrix. It can be run by the following command:

```
C:\WEITZMAN>weitzman matrix.txt (C:\WEITZMAN> - the system prompt)
```

The results will be saved in the text file called "w_output.txt". This file contains the table of distances by increasing order, the table of representatives (first column) and links (second column), amount of diversity and the location of the leaves in the tree.

Warning: every time you run the programme this file will be replaced automatically! On account of calculation time the program refuses to compute the diversity for the matrices with dimension larger than 30.

Example: w_output.txt

Table of distances, increasing order:

ISLR09	PLPU01	0.500000
DEBB01	CZPR01	1.000000
GBLR03	PLPU01	1.500000
DEBB01	GBLR03	2.000000
GBLR03	ISLR09	3.000000
GBLR03	CZPR01	3.500000
DEBB01	ISLR09	4.000000
PLPU01	CZPR01	5.000000
NOLR05	CZPR01	6.000000
DEBB01	PLPU01	7.000000
DEBB01	NOLR05	8.000000
NOLR05	PLPU01	10.000000
NOLR05	ISLR09	11.000000
ISLR09	CZPR01	12.000000
GBLR03	NOLR05	14.000000

Table of representatives and links:

Repr.	Link	Distances	Cumulated
ISLR09	[PLPU01]	0.500000	0.500000
CZPR01	[DEBB01]	1.000000	1.500000
ISLR09	[GBLR03]	3.000000	4.500000
CZPR01	[NOLR05]	6.000000	10.500000
ISLR09	[CZPR01]	12.000000	22.500000

Total diversity: 22.500000

Location of the leaves in the tree:

ISLR09 PLPU01 GBLR03 NOLR05 DEBB01 CZPR01

wmarg.exe

The programme "wmarg.exe" computes the marginal loss of diversity for each element of the set using the direct algorithm. It takes only one argument, the name of text file containing a square distance matrix. It can be run by the following command:

C:\WEITZMAN>wmarg matrix.txt

The results will be saved in the text file called "m_output.txt". This file contains the table of marginal losses of diversity. In the second line, second column you find the diversity of set $V(S)$. Below it the diversities of set without element i are listed $V(S-i)$. The third column represents the

marginal losses of diversity $dV(i) = V(S) - V(S-i)$, the fourth the marginal losses of diversity in percent.

Warning: every time you run the program this file will be replaced automatically. On account of calculation time the program refuses to compute the diversity for matrices with dimension larger than 30.

Example: m_output.txt

Name(i)	V(S-i)	dV(i)	dV(i) percent
V(S)	22.500000		
DEBB01	21.500000	1.000000	4.44
GBLR03	19.500000	3.000000	13.33
NOLR05	16.500000	6.000000	26.67
ISLR09	20.000000	2.500000	11.11
PLPU01	22.000000	0.500000	2.22
CZPR01	19.500000	3.000000	13.33

wgroup.exe

The programme "wgroup.exe" computes the marginal loss of diversity for predefined groups. The definition of groups must be done in the text file called "groups.txt" which must be placed in the same directory as the programme. If the programme does not find the file "groups.txt", the execution will be stopped.

Note: each element may be defined as a group in order to obtain its marginal loss of diversity.

The algorithm which will be used depends on the size of the distance matrix. For the groups larger than 30 elements the approximate method based on random simulation will be chosen. The programme takes two arguments: the first - the name of text file containing a square distance matrix, the second - the number of simulations for the approximate method. This number depends on the size of your matrix and of the reliability you want to obtain. For most trustworthy results use the maximal number of simulations which is possible on your machine regarding the calculation time (100.000 simulations take about an hour to obtain the marginal loss of diversity of one element for a matrix of 70 elements at PIII 500 MHz). The program can be run by the following command:

C:\WEITZMAN>wgroup matrix.txt 30000

The file "groups.txt" must be a simple text file and have the following form: in the first line - the number of groups, in the second line - the sizes of groups separated by spaces, starting with the third line - the names of elements belonging to each group separated by spaces. There are no restrictions to the composition of groups: the union of all the groups must not give the set; the groups can intersect.

Example: groups.txt

```
3
3 1 2
DEBB01 PLPU01 CZPR01
GBLR03
NOLR05 ISLR09
```

The results will be saved in the text file called "g_output.txt". This file contains the table of marginal losses of diversity. In the second line, second column you find the diversity of set V(S). Below it the diversities of groups are listed V(S-i). The third column represents the marginal losses of diversity dV(i), the fourth - the marginal losses of diversity in percent.

Warning: every time you run the programme this file will be replaced automatically!

Example: g_output.txt

Group i	V(S-i)	dV(i)	dV(i) percent
Set V(S)	22.500000		
1	17.000000	5.500000	24.44
2	19.500000	3.000000	13.33
3	9.500000	13.000000	57.78
