

Package ‘FD’

October 26, 2009

Type Package

Title Measuring functional diversity (FD) from multiple traits

Version 1.0-1

Date 2009-10-26

Author Etienne Laliberté

Maintainer Etienne Laliberté <etiennelaliberte@gmail.com>

Description FD is a package to compute different multidimensional FD indices. It implements a distance-based framework to measure FD that allows any number and type of functional traits, and can also consider species relative abundances.

License GPL-2

LazyLoad yes

LazyData yes

Depends ade4, ape, geometry, vegan

Encoding latin1

Repository CRAN

Date/Publication 2009-10-26 19:25:57

R topics documented:

FD-package	2
dbFD	3
dummy	10
fdisp	10
functcomp	12
gowdis	14
simul.dbFD	16
tussock	18

Index	20
--------------	-----------

FD-package

*Measure Functional Diversity from Multiple Traits***Description**

FD is a package to compute different multidimensional functional diversity (FD) indices. It implements a distance-based framework to measure FD that allows any number and type of functional traits, and can also consider species relative abundances.

Details

Package:	FD
Type:	Package
Version:	1.0-1
Date:	2009-10-22
License:	GPL-2
LazyLoad:	yes
LazyData:	yes

FD computes different multidimensional FD indices. To compute FD indices, a species-by-trait(s) matrix is required (or at least a species-by-species distance matrix). `gowdis` computes the Gower dissimilarity from different trait types (continuous, ordinal, nominal, or binary), and tolerates *NA*s. It can treat ordinal variables as described by Podani (1999), and can handle asymmetric binary variables and variable weights. `gowdis` is called by `dbFD`, the main function of **FD**.

`dbFD` uses principal coordinates analysis (PCoA) to return PCoA axes, which are then used as 'traits' to compute FD. `dbFD` computes several multidimensional FD indices, including the three indices of Villéger et al. (2008): functional richness (FRic), functional evenness (FEve), and functional divergence (FDiv). It also computes functional dispersion (FDis) (Laliberté and Legendre 2009), Rao's quadratic entropy (Q) (Botta-Dukát 2005), a posteriori functional group richness (FGR), and the community-level weighted means of trait values (CWM), an index of functional composition. Some of these indices can be weighted by species abundances. `dbFD` includes several options for flexibility.

Author(s)

Etienne Laliberté

Maintainer: Etienne Laliberté <etiennelaliberte@gmail.com> <http://www.elaliberte.info>**References**

Botta-Dukát, Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science* **16**:533-540.

Laliberté, E. and P. Legendre (2009) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **in press**.

Podani, J. (1999) Extending Gower's general coefficient of similarity to ordinal characters. *Taxon* **48**:331-340.

Villéger, S., N. W. H. Mason and D. Mouillot (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* **89**:2290-2301.

Examples

```
# examples with a dummy dataset

ex1 <- gowdis(dummy$trait)
ex1

ex2 <- functcomp(dummy$trait, dummy$abun)
ex2

ex3 <- dbFD(dummy$trait, dummy$abun)
ex3

# examples with real data from New Zealand short-tussock grasslands
# these examples may take a few seconds to a few minutes each to run

ex4 <- gowdis(tussock$trait)

ex5 <- functcomp(tussock$trait, tussock$abun)

# 'lingoes' correction used because 'sqrt' does not work in that case
ex6 <- dbFD(tussock$trait, tussock$abun, corr = "lingoes")

## Not run:
# ward clustering to compute FGR, cailliez correction
ex7 <- dbFD(tussock$trait, tussock$abun, corr = "cailliez",
  calc.FGR = TRUE, clust.type = "ward")
# choose 'g' for number of groups
# 6 groups seems to make good ecological sense
ex7

# however, calinski criterion in 'kmeans' suggests
# that 6 groups may not be optimal
ex8 <- dbFD(tussock$trait, tussock$abun, corr = "cailliez",
  calc.FGR = TRUE, clust.type = "kmeans", km.sup.gr = 10)
## End(Not run)
```

Description

dbFD implements a flexible distance-based framework to compute multidimensional functional diversity (FD) indices. dbFD returns the three FD indices of Villéger et al. (2008): functional

richness (FRic), functional evenness (FEve), and functional divergence (FDiv), as well functional dispersion (FDis; Laliberté and Legendre 2009), Rao's quadratic entropy (Q) (Botta-Dukát 2005), a posteriori functional group richness (FGR) (Petchey and Gaston 2006), and the community-level weighted means of trait values (CWM; e.g. Lavorel et al. 2008). Some of these FD indices consider species abundances. dbFD includes several options for flexibility.

Usage

```
dbFD(x, a, w, w.abun = TRUE, stand.x = TRUE,
     ord = c("podani", "metric"), asym.bin = NULL,
     corr = c("sqrt", "cailliez", "lingoes", "none"),
     calc.FRic = TRUE, m = "max", stand.FRic = TRUE,
     scale.RaoQ = FALSE, calc.FGR = FALSE, clust.type = "ward",
     km.inf.gr = 2, km.sup.gr = nrow(x) - 1, km.iter = 100,
     km.crit = c("calinski", "ssi"), calc.CWM = TRUE,
     CWM.type = c("dom", "all"), calc.FDiv = TRUE, dist.bin = 2,
     print.pco = FALSE, messages = TRUE)
```

Arguments

- | | |
|----------|--|
| x | matrix or data frame of functional traits. Traits can be numeric , ordered , or factor . Binary traits should be numeric and only contain 0 and 1. character traits will be converted to factor . NAs are tolerated.

x can also be a species-by-species distance matrix of class dist , in which case NAs are not allowed.

When there is only one trait, x can be also be a numeric vector, an ordered factor, or a unordered factor .

In all cases, species labels are required. |
| a | matrix containing the abundances of the species in x (or presence-absence, i.e. 0 or 1). Rows are sites and species are columns. Can be missing, in which case dbFD assumes that there is only one community with equal abundances of all species. NAs will be replaced by 0. The number of species (columns) in a must match the number of species (rows) in x. In addition, the species labels in a and x must be identical and in the same order. |
| w | vector listing the weights for the traits in x. Can be missing, in which case all traits have equal weights. |
| w.abun | logical; should FDis, Rao's Q, FEve, FDiv, and CWM be weighted by the relative abundances of the species? |
| stand.x | logical; if all traits are numeric , should they be standardized to mean 0 and unit variance? If not all traits are numeric , Gower's (1971) standardization by the range is automatically used; see gowdis for more details. |
| ord | character string specifying the method to be used for ordinal traits (i.e. ordered). "podani" refers to Eqs. 2a-b of Podani (1999), while "metric" refers to his Eq. 3. Can be abbreviated. See gowdis for more details. |
| asym.bin | vector listing the asymmetric binary variables in x. See gowdis for more details. |

<code>corr</code>	character string specifying the correction method to use when the species-by-species distance matrix cannot be represented in a Euclidean space. Options are "sqrt", "cailliez", "lingoes", or "none". Can be abbreviated. Default is "sqrt". See ‘details’ section.
<code>calc.FRic</code>	logical; should FRic be computed?
<code>m</code>	the number of PCoA axes to keep as ‘traits’ for calculating FRic (when FRic is measured as the convex hull volume) and FDiv. Options are: any integer > 1, "min" (maximum number of traits that allows the $s \geq 2^t$ condition to be met, where s is the number of species and t the number of traits), or "max" (maximum number of axes that allows the $s > t$ condition to be met). See ‘details’ section.
<code>stand.FRic</code>	logical; should FRic be standardized by the ‘global’ FRic that include all species, so that FRic is constrained between 0 and 1?
<code>scale.RaoQ</code>	logical; should Rao’s Q be scaled by its maximal value over all frequency distributions? See divc .
<code>calc.FGR</code>	logical; should FGR be computed?
<code>clust.type</code>	character string specifying the clustering method to be used to create the dendrogram of species for FGR. Options are "ward", "single", "complete", "average", "mcquitty", "median", "centroid", and "kmeans". For "kmeans", the argument <code>groups.kmeans</code> also applies. See hclust and cascadeKM for more details.
<code>km.inf.gr</code>	the number of groups for the partition with the smallest number of groups of the cascade (min). Only applies if <code>calc.FGR</code> is TRUE and <code>clust.type</code> is "kmeans". See cascadeKM for more details.
<code>km.sup.gr</code>	the number of groups for the partition with the largest number of groups of the cascade (max). Only applies if <code>calc.FGR</code> is TRUE and <code>clust.type</code> is "kmeans". See cascadeKM for more details.
<code>km.iter</code>	the number of random starting configurations for each value of K . Only applies if <code>calc.FGR</code> is TRUE and <code>clust.type</code> is "kmeans". See cascadeKM for more details.
<code>km.crit</code>	criterion used to select the best partition. The default value is "calinski" (Calinski-Harabasz 1974). The simple structure index "ssi" is also available. Only applies if <code>calc.FGR</code> is TRUE and <code>clust.type</code> is "kmeans". Can be abbreviated. See cascadeKM for more details.
<code>calc.CWM</code>	logical; should the community-level weighted means of trait values (CWM) be calculated? Can be abbreviated. See functcomp for more details.
<code>CWM.type</code>	character string indicating how nominal, binary and ordinal traits should be handled for CWM. See functcomp for more details.
<code>calc.FDiv</code>	logical; should FDiv be computed?
<code>dist.bin</code>	only applies when <code>x</code> is a single unordered factor , in which case <code>x</code> is coded using dummy variables. <code>dist.bin</code> is an integer between 1 and 10 specifying the appropriate distance measure for binary data. 2 (the default) refers to the simple matching coefficient (Sokal and Michener 1958). See dist.binary for the other options.

<code>print.pco</code>	logical; should the eigenvalues and PCoA axes be returned?
<code>messages</code>	logical; should warning messages be printed in the console?

Details

Typical usage is

```
dbFD(x, a, ...)
```

If `x` is a matrix or a data frame that contains only continuous traits, no `NA`s, and that no weights are specified (i.e. `w` is missing), a species-species Euclidean distance matrix is computed via `dist`. Otherwise, a Gower dissimilarity matrix is computed via `gowdis`. If `x` is a distance matrix, it is taken as is.

When `x` is a single trait, species with `NA`s are first excluded to avoid `NA`s in the distance matrix. If `x` is a single continuous trait (i.e. of class `numeric`), a species-species Euclidean distance matrix is computed via `dist`. If `x` is a single ordinal trait (i.e. of class `ordered`), `gowdis` is used and argument `ord` applies. If `x` is a single nominal trait (i.e. an unordered `factor`), the trait is converted to dummy variables and a distance matrix is computed via `dist.binary`, following argument `dist.bin`.

Once the species-species distance matrix is obtained, `dbFD` checks whether it is Euclidean. This is done via `is.euclid`. PCoA axes corresponding to negative eigenvalues are imaginary axes that cannot be represented in a Euclidean space, but simply ignoring these axes would lead to biased estimations of FD. Hence in `dbFD` one of four correction methods are used, following argument `corr`. `"sqrt"` simply takes the square root of the distances. However, this approach does not always work for all coefficients, in which case `dbFD` will stop and tell the user to select another correction method. `"cailliez"` refers to the approach described by Cailliez (1983) and is implemented via `cailliez`. `"lingoes"` refers to the approach described by Lingoes (1971) and is implemented via `lingoes`. `"none"` creates a distance matrix with only the positive eigenvalues of the Euclidean representation via `quasieuclid`. See Legendre and Legendre (1998) and Legendre and Anderson (1999) for more details on these corrections.

Principal coordinates analysis (PCoA) is then performed (via `dudi.pco`) on the *corrected* species-species distance matrix. The resulting PCoA axes are used as the new ‘traits’ to compute the three indices of Villéger et al. (2008): `FRic`, `FEve`, and `FDiv`. For `FEve`, there is no limit on the number of traits that can be used, so all PCoA axes are used. On the other hand, `FRic` and `FDiv` both rely on finding the minimum convex hull that includes all species (Villéger et al. 2008). This requires more species than traits. To circumvent this problem, `dbFD` takes only a subset of the PCoA axes as traits via argument `m`. This, however, comes at a cost of loss of information. The quality of the resulting reduced-space representation is returned by `qual.FRic`, which is computed as described by Legendre and Legendre (1998) and can be interpreted as a R^2 -like ratio.

In `dbFD`, `FRic` is generally measured as the convex hull volume, but when there is only one continuous trait it is measured as the range (or the range of the ranks for an ordinal trait). Conversely, when only nominal and ordinal traits are present, `FRic` is measured as the number of unique trait value combinations in a community. `FEve` and `FDiv`, but not `FRic`, can account for species relative abundances, as described by Villéger et al. (2008).

Functional dispersion (`FDIs`; Laliberté and Legendre 2009) is computed from the *uncorrected* species-species distance matrix via `fdisp`. Axes with negative eigenvalues are corrected following the approach of Anderson (2006). When all species have equal abundances (i.e. presence-absence data), `FDIs` is simply the average distance to the centroid (i.e. multivariate dispersion) as

originally described by Anderson (2006). Multivariate dispersion has been proposed as an index of beta diversity (Anderson et al. 2006). However, Laliberté and Legendre (2009) have extended it to a FD index. FDis can account for relative abundances by shifting the position of the centroid towards the most abundant species, and then computing a weighted average distance to this new centroid, using again the relative abundances as weights (Laliberté and Legendre 2009). FDis has no upper limit and requires at least two species to be computed. For communities composed of only one species, dbFD returns a FDis value of 0. FDis is by construction unaffected by species richness, it can be computed from any distance or dissimilarity measure (Anderson et al. 2006), it can handle any number and type of traits (including more traits than species), and it is not strongly influenced by outliers.

Rao's quadratic entropy (Q) is computed from the *uncorrected* species-species distance matrix via `divc`. See Botta-Dukát (2005) for details. Rao's Q is conceptually similar to FDis, and simulations (via `simul.dbFD`) have shown high positive correlations between the two indices (Laliberté and Legendre 2009). Still, one potential advantage of FDis over Rao's Q is that in the unweighted case (i.e. with presence-absence data), it opens possibilities for formal statistical tests for differences in FD between two or more communities through a distance-based test for homogeneity of multivariate dispersions (Anderson 2006); see `betadisper` for more details.

Functional group richness (FGR) is based on the classification of the species by the user from visual inspection of a dendrogram. Method "kmeans" is also available by calling `cascadeKM`. In that case, the Calinski-Harabasz (1974) criterion or the simple structure index (SSI) can be used to estimate the number of functional groups; see `cascadeKM` for more details. FGR returns the number of functional groups per community, as well as the abundance of each group in each community.

The community-level means of trait values (CWM) is an index of functional composition (Loreau et al. 2008), and is computed via `functcomp`. Species with NAs for a given trait are excluded for that trait.

Value

<code>nbsp</code>	vector listing the number of species in each community
<code>sing.sp</code>	vector listing the number of functionally singular species in each community. If all species are functionally different, <code>sing.sp</code> will be identical to <code>nbsp</code> .
<code>FRic</code>	vector listing the FRic of each community
<code>qual.FRic</code>	quality of the reduced-space representation required to compute FRic and FDiv.
<code>FEve</code>	vector listing the FEve of each community
<code>FDiv</code>	vector listing the FDiv of each community. Only returned if <code>calc.FDiv</code> is TRUE.
<code>FDis</code>	vector listing the FDis of each community
<code>RaoQ</code>	vector listing the Rao's quadratic entropy (Q) of each community
<code>FGR</code>	vector listing the FGR of each community. Only returned if <code>calc.FGR</code> is TRUE.
<code>spfgr</code>	vector specifying functional group membership for each species. Only returned if <code>calc.FGR</code> is TRUE.
<code>gr.abun</code>	matrix containing the abundances of each functional group in each community. Only returned if <code>calc.FGR</code> is TRUE.

CWM	data frame containing the community-level weighted trait means (CWM). Only returned if <code>calc.CWM</code> is TRUE.
<code>x.values</code>	eigenvalues from the PCoA. Only returned if <code>print.pco</code> is TRUE.
<code>x.axes</code>	PCoA axes. Only returned if <code>print.pco</code> is TRUE.

Note

dbFD borrows code from the `F_RED` function of Villéger et al. (2008).

Author(s)

Etienne Laliberté (etiennelaliberte@gmail.com) <http://www.elaliberte.info>

References

- Anderson, M. J. (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62**:245-253.
- Anderson, M. J., K. E. Ellingsen and B. H. McArdle (2006) Multivariate dispersion as a measure of beta diversity. *Ecology Letters* **9**:683-693.
- Botta-Dukát, Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science* **16**:533-540.
- Cailliez, F. (1983) The analytical solution of the additive constant problem. *Psychometrika* **48**:305-310.
- Calinski, T. and J. Harabasz (1974) A dendrite method for cluster analysis. *Communications in Statistics* **3**:1-27.
- Gower, J. C. (1971) A general coefficient of similarity and some of its properties. *Biometrics* **27**:857-871.
- Laliberté, E. and P. Legendre (2009) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **in press**.
- Lavorel, S., K. Grigulis, S. McIntyre, N. S. G. Williams, D. Garden, J. Dorrough, S. Berman, F. Quétier, A. Thebault and A. Bonis (2008) Assessing functional diversity in the field - methodology matters! *Functional Ecology* **22**:134-147.
- Legendre, P. and M. J. Anderson (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* **69**:1-24.
- Legendre, P. and L. Legendre (1998) *Numerical Ecology*. 2nd English edition. Amsterdam: Elsevier.
- Lingoes, J. C. (1971) Some boundary conditions for a monotone analysis of symmetric matrices. *Psychometrika* **36**:195-203.
- Podani, J. (1999) Extending Gower's general coefficient of similarity to ordinal characters. *Taxon* **48**:331-340.
- Sokal, R. R. and C. D. Michener (1958) A statistical method for evaluating systematic relationships. *The University of Kansas Scientific Bulletin* **38**:1409-1438.
- Villéger, S., N. W. H. Mason and D. Mouillot (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* **89**:2290-2301.

See Also

[gowdis](#), [functcomp](#), [fdisp](#), [simul.dbFD](#), [divc](#), [treedive](#), [betadisper](#)

Examples

```
# mixed trait types, NA's
ex1 <- dbFD(dummy$trait, dummy$abun)
ex1

# add variable weights
# 'cailliez' correction is used because 'sqrt' does not work
w<-c(1, 5, 3, 2, 5, 2, 6, 1)
ex2 <- dbFD(dummy$trait, dummy$abun, w, corr="cailliez")

# if 'x' is a distance matrix
trait.d <- gowdis(dummy$trait)
ex3 <- dbFD(trait.d, dummy$abun)
ex3

# one numeric trait, one NA
num1 <- dummy$trait[,1] ; names(num1) <- rownames(dummy$trait)
ex4 <- dbFD(num1, dummy$abun)
ex4

# one ordered trait, one NA
ord1 <- dummy$trait[,5] ; names(ord1) <- rownames(dummy$trait)
ex5 <- dbFD(ord1, dummy$abun)
ex5

# one nominal trait, one NA
fac1 <- dummy$trait[,3] ; names(fac1) <- rownames(dummy$trait)
ex6 <- dbFD(fac1, dummy$abun)
ex6

# example with real data from New Zealand short-tussock grasslands
# 'lingoes' correction used because 'sqrt' does not work in that case
ex7 <- dbFD(tussock$trait, tussock$abun, corr = "lingoes")

## Not run:
# calc.FGR = T, 'ward'
ex7 <- dbFD(dummy$trait, dummy$abun, calc.FGR = T)
ex7

# calc.FGR = T, 'kmeans'
ex8 <- dbFD(dummy$trait, dummy$abun, calc.FGR = T,
  clust.type = "kmeans")
ex8

# ward clustering to compute FGR
ex9 <- dbFD(tussock$trait, tussock$abun,
  corr = "cailliez", calc.FGR = TRUE, clust.type = "ward")
# choose 'g' for number of groups
```

```
# 6 groups seems to make good ecological sense
ex9

# however, calinski criterion in 'kmeans' suggests
# that 6 groups may not be optimal
ex10 <- dbFD(tussock$trait, tussock$abun, corr = "cailliez",
calc.FGR = TRUE, clust.type = "kmeans", km.sup.gr = 10)
## End(Not run)
```

dummy

Dummy Functional Trait Dataset

Description

A dummy dataset containing 8 species and 8 functional traits (2 continuous, 2 nominal, 2 ordinal, and 2 binary), with some missing values. Also includes a matrix of species abundances from 10 communities.

Usage

```
dummy
```

Format

trait data frame of 8 functional traits on 8 species

abun matrix of abundances of the 8 species from 10 communities

Source

Etienne Laliberté (etiennelaliberte@gmail.com)

fdisp

Functional Dispersion

Description

`fdisp` measures the functional dispersion (FDis) of a set of communities, as described by Laliberté and Legendre (2009).

Usage

```
fdisp(d, a, tol = 1e-07)
```

Arguments

<code>d</code>	a species-by- species distance matrix computed from functional traits, such as that returned by <code>dist</code> or <code>gowdis</code> . <code>NA</code> s are not allowed.
<code>a</code>	matrix containing the abundances of the species in <code>d</code> (or presence-absence, i.e. 0 or 1). Rows are sites and species are columns. Can be missing, in which case <code>fdisp</code> assumes that there is only one community with equal abundances of all species. <code>NA</code> s will be replaced by 0. The number of species (columns) in <code>a</code> must match the number of species in <code>d</code> . In addition, the species labels in <code>a</code> and <code>d</code> must be identical and in the same order.
<code>tol</code>	tolerance threshold to test whether the distance matrix is Euclidean : an eigenvalue is considered positive if it is larger than $-tol * \lambda_1$, where λ_1 is the largest eigenvalue.

Details

`fdisp` computes, for a set of communities, the average distance of individual objects (species) in PCoA space from any distance or dissimilarity measure, as described by Anderson (2006). The average distance to the centroid is a measure of multivariate dispersion and as been suggested as an index of beta diversity (Anderson et al. 2006). However, in `fdisp` both the centroid and the average distance to this centroid can be weighted by individual objects. In other words, `fdisp` returns the weighted average distance to the weighted centroid. This was suggested so that multivariate dispersion could be used as a multidimensional functional diversity (FD) index that can be weighted by species abundances. This FD index has been called functional dispersion (FDis) and is described by Laliberté and Legendre (2009).

In sum, FDis can account for relative abundances by shifting the position of the centroid towards the most abundant species, and then computing a weighted average distance to this new centroid, using again the relative abundances as weights (Laliberté and Legendre 2009). FDis has no upper limit and requires at least two species to be computed. For communities composed of only one species, `dbFD` returns a FDis value of 0. FDis is by construction unaffected by species richness, it can be computed from any distance or dissimilarity measure (Anderson et al. 2006), it can handle any number and type of traits (including more traits than species), and it is not strongly influenced by outliers.

FDis is conceptually similar to Rao's quadratic entropy Q (Botta-Dukát 2005), and simulations (via `simul.dbFD`) have shown high positive correlations between the two indices (Laliberté and Legendre 2009). Still, one potential advantage of FDis over Rao's Q is that in the unweighted case (i.e. with presence-absence data), it opens possibilities for formal statistical tests for differences in FD between two or more communities through a distance-based test for homogeneity of multivariate dispersions (Anderson 2006); see `betadisper` for more details.

Corrections for PCoA axes corresponding to negative eigenvalues are applied following Anderson (2006); see also `betadisper` for more details on these corrections.

Value

<code>FDis</code>	vector listing the FDis of each community
<code>eig</code>	vector listing the eigenvalues of the PCoA
<code>vectors</code>	matrix containing the PCoA axes

Note

`fdisp` is implemented in `dbFD` and is used to compute the functional dispersion (FD_{is}) index.

Author(s)

Etienne Laliberté (etiennelaliberte@gmail.com) <http://www.elaliberte.info>

References

- Anderson, M. J. (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62**:245-253.
- Anderson, M. J., K. E. Ellingsen and B. H. McArdle (2006) Multivariate dispersion as a measure of beta diversity. *Ecology Letters* **9**:683-693.
- Botta-Dukát, Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science* **16**:533-540.
- Laliberté, E. and P. Legendre (2009) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **in press**.

See Also

`dbFD` for computing multidimensional FD indices and `betadisper` from which `fdisp` borrows some code.

Examples

```
# dummy dataset
dummy.dist <- gowdis(dummy$trait)
ex1 <- fdisp(dummy.dist, dummy$abun)
ex1

# example with real data from New Zealand short-tussock grasslands
ex2 <- fdisp(gowdis(tussock$trait), tussock$abun)
ex2
```

functcomp

Functional Composition

Description

`functcomp` returns the functional composition of a set of communities, as measured by the community-level weighted means of trait values (CWM; e.g. Lavorel et al. 2008).

Usage

```
functcomp(x, a, CWM.type = c("dom", "all"))
```

Arguments

<code>x</code>	matrix or data frame containing the functional traits. Traits can be <code>numeric</code> , <code>ordered</code> , or <code>factor</code> . Binary traits should be <code>numeric</code> and only contain 0 and 1. <code>character</code> traits will be converted to <code>factor</code> . For a given trait, species with <code>NA</code> are excluded.
<code>a</code>	matrix containing the abundances of the species in <code>x</code> (or presence-absence, i.e. 0 or 1). Rows are sites and species are columns. The number of species (columns) in <code>a</code> must match the number of species (rows) in <code>x</code> . In addition, the species labels in <code>a</code> and <code>x</code> must be identical and in the same order. <code>NA</code> s will be replaced by 0.
<code>CWM.type</code>	character string indicating how nominal, binary and ordinal traits should be handled. See ‘details’.

Details

`functcomp` computes the community-level weighted means of trait values for a set of communities (i.e. sites). For a continuous trait, CWM is the mean trait value of all species present in the community (after excluding species with `NA`s), weighted by their relative abundances.

For ordinal, nominal and binary traits, either the dominant class is returned (when `CWM.type` is "dom"), or the abundance of each individual class is returned (when `CWM.type` is "all").

When `CWM.type = "dom"`, if the maximum abundance value is shared between two or more classes, then one of these classes is randomly selected for CWM. Because species with `NA`s for a given trait are excluded for that trait, it is possible that when `CWM.type` is set to "all", the sum of the abundances of all classes for a given ordinal/nominal/binary trait does not equal the sum of the species abundances. Thus, it is definitely not recommended to have `NA`s for very abundant species, as this will lead to biased estimates of functional composition.

Value

a data frame containing the CWM values of each trait for each community.

Note

`functcomp` is implemented in `dbFD` and will be returned if `calc.CWM` is `TRUE`.

Author(s)

Etienne Laliberté (etiennelaliberte@gmail.com) <http://www.elaliberte.info>

References

Lavorel, S., K. Grigulis, S. McIntyre, N. S. G. Williams, D. Garden, J. Dorrough, S. Berman, F. Quétier, A. Thebault and A. Bonis (2008) Assessing functional diversity in the field - methodology matters! *Functional Ecology* **22**:134-147.

See Also

`dbFD` for measuring distance-based multidimensional functional diversity indices, including CWM.

Examples

```
# for ordinal, nominal and binary variables
# returns only the most frequent class
ex1 <- functcomp(dummy$trait, dummy$abun)
ex1

# returns the frequencies of each class
ex2 <- functcomp(dummy$trait, dummy$abun, CWM.type = "all")
ex2

# example with real data from New Zealand short-tussock grasslands
ex3 <- functcomp(tussock$trait, tussock$abun)
ex3
```

gowdis

Gower Dissimilarity

Description

`gowdis` measures the Gower (1971) dissimilarity for mixed variables, including asymmetric binary variables. Variable weights can be specified. `gowdis` implements Podani's (1999) extension to ordinal variables.

Usage

```
gowdis(x, w, asym.bin = NULL, ord = c("podani", "metric"))
```

Arguments

<code>x</code>	matrix or data frame containing the variables. Variables can be <code>numeric</code> , <code>ordered</code> , or <code>factor</code> . Symmetric or asymmetric binary variables should be <code>numeric</code> and only contain 0 and 1. <code>character</code> variables will be converted to <code>factor</code> . NAs are tolerated.
<code>w</code>	vector listing the weights for the variables in <code>x</code> . Can be missing, in which case all variables have equal weights.
<code>asym.bin</code>	vector listing the asymmetric binary variables in <code>x</code> .
<code>ord</code>	character string specifying the method to be used for ordinal variables (i.e. <code>ordered</code>). "podani" refers to Eqs. 2a-b of Podani (1999), while "metric" refers to his Eq. 3 (see 'details'). Can be abbreviated.

Details

`gowdis` computes the Gower (1971) similarity coefficient exactly as described by Podani (1999), then converts it to a dissimilarity coefficient by using $D = 1 - S$. It integrates variable weights as described by Legendre and Legendre (1998).

Let $\mathbf{X} = \{x_{ij}\}$ be a matrix containing n objects (rows) and m columns (variables). The similarity G_{jk} between objects j and k is computed as

$$G_{jk} = \frac{\sum_{i=1}^n w_{ijk} s_{ijk}}{\sum_{i=1}^n w_{ijk}}$$

,
 where w_{ijk} is the weight of variable i for the j - k pair, and s_{ijk} is the partial similarity of variable i for the j - k pair,

and where $w_{ijk} = 0$ if objects j and k cannot be compared because x_{ij} or x_{ik} is unknown (i.e. NA).

For binary variables, $s_{ijk} = 0$ if $x_{ij} \neq x_{ik}$, and $s_{ijk} = 1$ if $x_{ij} = x_{ik} = 1$ or if $x_{ij} = x_{ik} = 0$.

For asymmetric binary variables, same as above except that $w_{ijk} = 0$ if $x_{ij} = x_{ik} = 0$.

For nominal variables, $s_{ijk} = 0$ if $x_{ij} \neq x_{ik}$ and $s_{ijk} = 1$ if $x_{ij} = x_{ik}$.

For continuous variables,

$$s_{ijk} = 1 - \frac{|x_{ij} - x_{ik}|}{x_{i.max} - x_{i.min}}$$

where $x_{i.max}$ and $x_{i.min}$ are the maximum and minimum values of variable i , respectively.

For ordinal variables, all x_{ij} are replaced by their ranks r_{ij} determined over all objects (such that ties are also considered), and then

if ord = "podani"

$s_{ijk} = 1$ if $r_{ij} = r_{ik}$, otherwise

$$s_{ijk} = 1 - \frac{|r_{ij} - r_{ik}| - (T_{ij} - 1)/2 - (T_{ik} - 1)/2}{r_{i.max} - r_{i.min} - (T_{i.max} - 1)/2 - (T_{i.min} - 1)/2}$$

where T_{ij} is the number of objects which have the same rank score for variable i as object j (including j itself), T_{ik} is the number of objects which have the same rank score for variable i as object k (including k itself), $r_{i.max}$ and $r_{i.min}$ are the maximum and minimum ranks for variable i , respectively, $T_{i.max}$ is the number of objects with the maximum rank, and $T_{i.min}$ is the number of objects with the minimum rank.

if ord = "metric"

$$s_{ijk} = 1 - \frac{|r_{ij} - r_{ik}|}{r_{i.max} - r_{i.min}}$$

Value

an object of class `dist` with the following attributes: Labels, Types (the variable types, where 'C' is continuous/numeric, 'O' is ordinal, 'B' is symmetric binary, 'A' is asymmetric binary, and 'N' is nominal), Size, Metric.

Author(s)

Etienne Laliberté (etiennelaliberte@gmail.com) <http://www.elaliberte.info>, with some help from Philippe Casgrain for the C interface.

References

- Gower, J. C. (1971) A general coefficient of similarity and some of its properties. *Biometrics* **27**:857-871.
- Legendre, P. and L. Legendre (1998) *Numerical Ecology*. 2nd English edition. Amsterdam: Elsevier.
- Podani, J. (1999) Extending Gower's general coefficient of similarity to ordinal characters. *Taxon* **48**:331-340.

See Also

[daisy](#) is similar but less flexible, since it does not include variable weights and does not treat ordinal variables as described by Podani (1999).

Examples

```
ex1 <- gowdis(dummy$trait)
ex1

# check attributes
attributes(ex1)

# to include weights
w <- c(4, 3, 5, 1, 2, 8, 3, 6)
ex2 <- gowdis(dummy$trait, w)
ex2

# variable 7 as asymmetric binary
ex3 <- gowdis(dummy$trait, asym.bin = 7)
ex3

# example with trait data from New Zealand vascular plant species
# takes a few seconds
ex4 <- gowdis(tussock$trait)
```

simul.dbFD

Simulations to Explore Relationships Between Functional Diversity Indices

Description

simul.dbFD generates artificial communities of species with artificial functional traits. Different functional diversity (FD) indices are computed from these communities using [dbFD](#) to explore their inter-relationships.

Usage

```
simul.dbFD(s = c(5, 10, 15, 20, 25, 30, 35, 40), t = 3,
           r = 10, p = 100, tr.method = c("unif", "norm", "lnorm"),
           abun.method = c("lnorm", "norm", "unif"), w.abun = TRUE)
```

Arguments

<code>s</code>	vector listing the different levels of species richness used in the simulations
<code>t</code>	number of traits
<code>r</code>	number of replicates per species richness level
<code>p</code>	number of species in the common species pool
<code>tr.method</code>	character string indicating the sampling distribution for the traits. "unif" is a uniform distribution, "norm" is a normal distribution, and "lnorm" is a lognormal distribution.
<code>abun.method</code>	character string indicating the sampling distribution for the species abundances. Same as for <code>tr.method</code> .
<code>w.abun</code>	logical; should FDis, FEve, FDiv, and Rao's quadratic entropy (Q) be weighted by species abundances?

Value

A list containing the following elements:

<code>results</code>	data frame containing the results of the simulations
<code>traits</code>	matrix containing the traits
<code>abun</code>	matrix containing the abundances
<code>abun.gamma</code>	species abundances from the pooled set of communities
<code>FDis.gamma</code>	FDis of the pooled set of communities
<code>FDis.mean</code>	mean FDis from all communities

`FDis.gamma` and `FDis.mean` can be used to explore the set concavity criterion (Ricotta 2005) for FDis.

A graph plotting the results of the simulations is also returned.

Warning

The simulations performed by `simul.dbFD` can take several hours if `length(s)` and/or `r` is large. Run a test with the default parameters first.

Author(s)

Etienne Laliberté (etiennelaliberte@gmail.com) <http://www.elaliberte.info>

References

- Laliberté, E. and P. Legendre (2009) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **in press**.
- Ricotta, C. (2005) A note on functional diversity measures. *Basic and Applied Ecology* **6**:479-486.

See Also

[dbFD](#), the function called in `simul.dbFD`

Examples

```
# this should take just a few minutes
## Not run:
ex1 <- simul.dbFD(s = c(10, 20, 30, 40, 50), r = 5)
ex1
## End(Not run)
```

tussock

Functional Composition of Short-Tussock Grasslands

Description

tussock contains data on 16 functional traits measured on 53 vascular plant species from New Zealand short-tussock grasslands. It also contains the relative abundances (percent cover) of these 53 species from 30 8x50-m plots.

Usage

```
tussock
```

Format

tussock is a list of 2 components:

trait data frame of 16 functional traits measured on 53 plant species: growth form (sensu Cornelissen et al. 2003), reproductive plant height (m), leaf dry matter content (mg g^{-1}), leaf nitrogen concentration (mg g^{-1}), leaf phosphorous concentration (mg g^{-1}), leaf sulphur concentration (mg g^{-1}), specific leaf area ($\text{m}^2 \text{kg}^{-1}$), nutrient uptake strategy (sensu Cornelissen et al. 2003), Raunkiaer life form, clonality, leaf size (mm^2), primary dispersal mode, seed mass (mg), resprouting capacity, pollination syndrome, and lifespan (an ordinal variable stored as `ordered`).

abun matrix containing the relative abundances (percent cover) of the 53 species in 30 plots

Details

The functional traits were measured using standardized methodologies (Cornelissen et al. 2003). Each of the 30 experimental plots from which species cover was estimated is 8x50 m. Relative abundances of all vascular plant species were estimated in November 2007. To do so, 20 1x1-m quadrats per plot were randomly positioned along two longitudinal transects and cover of each species was estimated using a modified Braun-Blanquet scale. This data was pooled at the plot scale to yield the percent cover data.

Source

Etienne Laliberté (etiennelaliberte@gmail.com)

<http://www.elaliberte.info>

References

- Cornelissen, J. H. C., S. Lavorel, E. Garnier, S. Diaz, N. Buchmann, D. E. Gurvich, P. B. Reich, H. ter Steege, H. D. Morgan, M. G. A. van der Heijden, J. G. Pausas and H. Poorter. (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany* **51**:335-380.
- Laliberté, E., Norton, D. A. and D. Scott. (2008) Impacts of rangeland development on plant functional diversity, ecosystem processes and services, and resilience. *Global Land Project (GLP) Newsletter* **4**:4-6.
- Scott, D. (1999) Sustainability of New Zealand high-country pastures under contrasting development inputs. 1. Site, and shoot nutrients. *New Zealand Journal of Agricultural Research* **42**:365-383.

Index

- *Topic **datagen**
 - simul.dbFD, 16
- *Topic **datasets**
 - dummy, 10
 - tussock, 18
- *Topic **multivariate**
 - dbFD, 3
 - fdisp, 10
 - functcomp, 12
 - gowdis, 14
- *Topic **package**
 - FD-package, 2

betadisper, 7, 9, 11, 12

cailliez, 6

cascadeKM, 5, 7

character, 4, 13, 14

daisy, 16

dbFD, 2, 3, 12, 13, 16, 17

dist, 4, 6, 11, 15

dist.binary, 5, 6

divc, 5, 7, 9

dudi.pco, 6

dummy, 10

factor, 4–6, 13, 14

FD (*FD-package*), 2

FD-package, 2

fdisp, 6, 9, 10

functcomp, 5, 7, 9, 12

gowdis, 2, 4, 6, 9, 11, 14

hclust, 5

is.euclid, 6

lingoes, 6

NA, 2, 4, 6, 7, 11, 13–15

numeric, 4, 6, 13, 14

ordered, 4, 6, 13, 14, 18

quasieuclid, 6

simul.dbFD, 7, 9, 11, 16

treedive, 9

tussock, 18