

Appendix 1

Functional diversity FD_Q

Calculation of functional diversity is based on a functional trait matrix X containing the values x_{st} of trait t of species s ($s=1, \dots, S$, $t=1 \dots T$),

$$X = (x_1, \dots, x_t) = \begin{pmatrix} x_{11} & \dots & x_{1T} \\ \vdots & \ddots & \vdots \\ x_{S1} & \dots & x_{ST} \end{pmatrix}.$$

Rao's quadratic diversity FD_Q is defined as

$$FD_Q = \sum_{i=1}^S \sum_{j=1}^S p_i p_j d_{ij}$$

where S is the number of species in the community, p_i and p_j are the relative abundances of species i and j , and d_{ij} describes the functional dissimilarity between species i and j . There exists a vast number of different dissimilarity measures used in multivariate statistical analysis which can be derived from the trait matrix (Legendre and Legendre 1998). Whatever dissimilarity measure is chosen, functional diversity FD_Q can be interpreted as the expected dissimilarity of two individuals randomly chosen from the community. We restrict attention to one particular dissimilarity measure, the squared Euclidean distance in the multidimensional trait space, for reasons which will become apparent. Firstly, it has a convenient property which will become important in the applications below: The resulting diversity index

$$FD_Q = \sum_{i=1}^S \sum_{j=1}^S p_i p_j \left(\sum_{t=1}^T (x_{it} - x_{jt})^2 \right) = \sum_{t=1}^T \left(\sum_{i=1}^S \sum_{j=1}^S p_i p_j (x_{it} - x_{jt})^2 \right)$$

is additive with respect to traits. Thus the contribution of different functional traits to the functional diversity of a community can easily be partitioned. Secondly, it can be shown (Rao 1986) that

$$\sum_{i=1}^S \sum_{j=1}^S p_i p_j (x_{it} - x_{jt})^2 = 2\text{Var}(x_t)$$

thus this choice of dissimilarity measure relates FD_Q to variances of individual traits, a commonly used measure of variability.

Estimating relative weights of functional traits

If the calculation of FD_Q is based on multiple functional traits, the different traits are implicitly considered equally important. However, it is very unlikely that all traits contribute in a comparable manner to the functional diversity of a community and the assessment of the relative importance of particular traits is often an important scientific problem. Petchey et al. (2004) first used their functional diversity index FD to identify the functionally important traits by comparing the explanatory power of different combinations of candidate traits. Their qualitative inclusion/exclusion of traits can be extended quantitatively based on FD_Q in the following way: each candidate trait is given an a priori unknown weight α_t which will be estimated from the data.

Incorporation of weights into the trait matrix

$$X^* = \begin{pmatrix} x_{11}^* & \cdots & x_{1T}^* \\ \vdots & \ddots & \vdots \\ x_{S1}^* & \cdots & x_{ST}^* \end{pmatrix} = \begin{pmatrix} \alpha_1 x_{11} & \cdots & \alpha_T x_{1T} \\ \vdots & \ddots & \vdots \\ \alpha_1 x_{S1} & \cdots & \alpha_T x_{ST} \end{pmatrix}$$

results in a trait-weighted version of the functional diversity measure

$$\begin{aligned} FD_Q &= \sum_{i=1}^S \sum_{j=1}^S p_i p_j \left(\sum_{t=1}^T (\alpha_t x_{it} - \alpha_t x_{jt})^2 \right) = \sum_{t=1}^T \alpha_t^2 \left(\sum_{i=1}^S \sum_{j=1}^S p_i p_j (x_{it} - x_{jt})^2 \right) \\ &= 2 \sum_{t=1}^T \alpha_t^2 \text{Var}(x_t) \end{aligned}$$

Using functional diversity as an explanatory variable for the aboveground community biomass of semi-natural grasslands requires calculation of FD_Q for all communities considered:

$$FD_{Q_m} = \sum_{t=1}^T \alpha_t^2 \sum_{i \in I_m} \sum_{j \in I_m} p_i p_j (x_{it} - x_{jt})^2,$$

where $m = 1, \dots, M$ indicate the different communities and $I_m \subseteq \{1, 2, \dots, S\}$ is an index set containing all the species from the total species pool present in a particular community m . The simplest possible model; a linear regression model with FD_Q as the single linear predictor for aboveground biomass, can therefore be written as

$$\begin{aligned} y_m &= \beta_0 + \beta_1 FD_{Q_m}, \quad m = 1, \dots, M \\ &= \beta_0 + \beta_1 \sum_{t=1}^T \alpha_t^2 \sum_{i \in I_m} \sum_{j \in I_m} p_i p_j (x_{it} - x_{jt})^2 \\ &= \beta_0 + \sum_{t=1}^T (\beta_1 \alpha_t^2) \sum_{i \in I_m} \sum_{j \in I_m} p_i p_j (x_{it} - x_{jt})^2 \\ &= \beta_0 + \sum_{t=1}^T \gamma_t z_{mt}, \end{aligned}$$

which is a multiple linear regression model with the response variable y_m being the aboveground biomass in community m and T predictor variables z_{mt} , $t = 1, \dots, T$, describing diversity of community m with respect to trait t . The particular form of the regression coefficients $\gamma_t = \beta_1 \alpha_t^2$ implies an additional constraint – all γ_t must have the same sign. Estimation can therefore rely on least squares procedures with box constraints (all γ_t positive or negative, respectively) available in standard statistical packages. If the estimation procedure results in zero values for some γ_t , inclusion of these traits into the calculation of functional diversity would not increase its predictive power with respect to the response variable in question.

Because $T+1$ parameters were used to define the T regression coefficients γ_t , one ancillary condition must be specified to make the parameters identifiable. Fixing the largest α_t to 1, thus defining

$$\begin{aligned} \beta_1 &= \max_{t=1, \dots, T} \gamma_t \\ \alpha_t^2 &= \frac{\gamma_t}{\max_{t=1, \dots, T} \gamma_t} = \frac{\gamma_t}{\beta_1}, \end{aligned}$$

emphasizes the interpretation of α_t as the relative importance of diversity with respect to trait t for aboveground community biomass of the studied grassland communities (This interpretation naturally requires that the traits are standardized to unit variance beforehand). Using this multiple linear regression approach we can therefore not only assess whether functional diversity is an important predictor for aboveground community biomass but also estimate the relative contribution of different traits to functional diversity of plant communities with respect to aboveground biomass as an important component of ecosystem functioning.

Extension of the pure diversity model

The approach of estimating relative weights for different functional traits to maximize explanatory power of FD_Q can easily be extended to models including other predictor variables, e.g. abiotic characteristics of sites or abundance-weighted aggregated traits of plant communities. The full regression model is

$$y_m = \beta_0 + \sum_{i=1}^I \lambda_i a_{mi} + \sum_{t=1}^T \mu_t b_{mt} + \sum_{t=1}^T \gamma_t z_{mt},$$

where a_{mi} , $I = 1, \dots, I$, denote the abiotic variables measured at site m with corresponding regression coefficients λ_i and b_{mt} are the abundance-weighted aggregated trait values at site m with corresponding regression coefficients μ_t .

References

- Legendre, P. and Legendre, L. 1998. Numerical ecology. – Elsevier.
- Petchey, O. L. et al. 2004. How do different measures of functional diversity perform? – Ecology 85: 847–857.
- Rao, C. R. 1986. Rao's axiomatization of diversity measures. – In: Kotz, S. and Johnson, N. L. (eds), Encyclopedia of statistical sciences. Wiley, pp. 614–617.