Draw pattern of alpha and gamma diversity along latitude using Gentry's forest plots

David Zelený & Po-Ju Ke

Introduction

Alwyn Howard Gentry (1945-1993) was an American botanist and taxonomist, who developed a sampling design for a rapid inventory of diversity in species-rich tropical forests. In a relatively homogeneous forest stand, he placed **ten contiguous transects** 2 m × 50 m, with the long and narrow shape of the transect allowing fast census (Fig. 1). Within each transect, he measured diameter in breast height (DBH) of each individual of woody species and lianas with DBH at least 2.5 cm.

After Gentry's tragic death (in the flight accident while on the way for a field trip in Western Ecuador), the dataset was published in printed version (Phillips & Miller 2002) and also made available online.
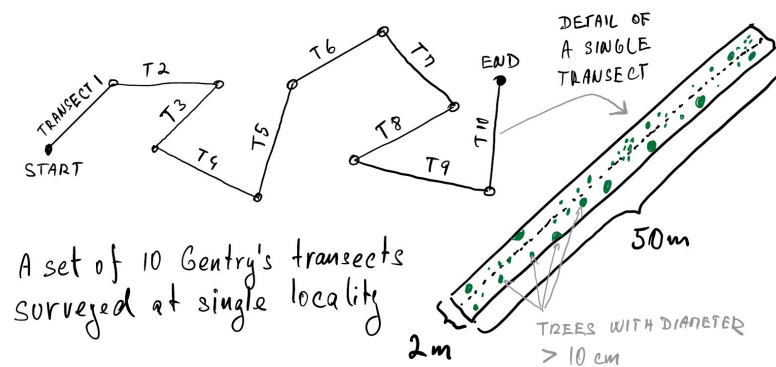

Figure 1: Design of Gentry's forest transects sampled on one of the localities. The detail of a single transect is on the right, with trees indicated by filled green circles. Ten transects per locality have been surveyed. (Schema by D. Zelený.)

Nathan Kraft and his colleagues took this dataset and in 2011 published the paper in *Science* (here), where they showed (among others) also the pattern of alpha and gamma diversity along the latitudinal gradient (Fig. 2 below, their Fig. 1A). It shows the relationship between mean alpha diversity (mean number of species in each transect within the locality) and gamma diversity (total number of species in locality) along the latitudinal gradient. The most diverse are tropics (low latitude), and diversity gets lower when moving north or south of the Equator (toward higher latitudes).

Please, recreate this figure, using the data which we independently downloaded from the original website and prepared for you to use.
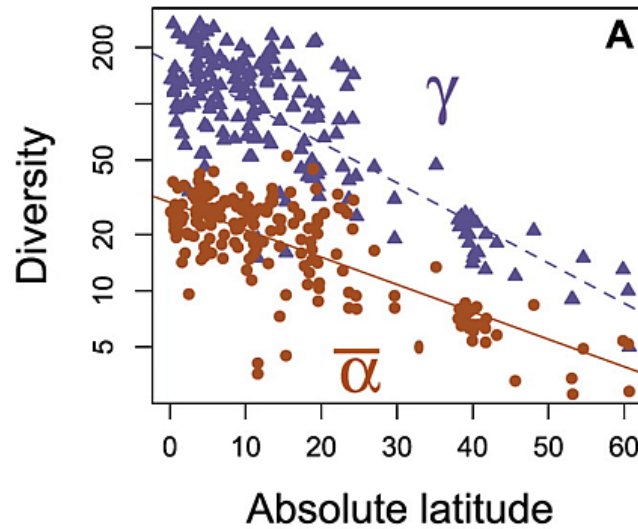
Figure 2: Kraft et al. (2011) (their Fig. 1A)

What to do

1. To do so, first load the data into R (the data are ready to be uploaded to R). You need (1) species composition of individual transects on each of 197 localities and (2) latitude of each locality. Import the following two files: {https://github.com/zdealveindy/recol/blob/main/data/gentry197.RData?raw=true} and {https://raw.githubusercontent.com/zdealveindy/recol/main/data/gentry.coord.txt}. The first file is an R object `gentry197` (list of data frames with species composition data) saved as a file `gentry197.RData`. You need to use the function `load` to load these data - and you should do it directly from the link above (without need to store the file on your computer). The second file is a matrix of coordinates saved in the tab-delimited format as *.txt file.

2. After importing into R, the object `gentry197` will appear in your working space. It is a list of 197 components, where each component is a data frame from one locality. Each locality's data frame has ten rows (one row = one transect) and many columns (one column = one species). Each cell of the data frame represents *a number of individuals* of given species in the given transect (information about DBH of individual trees is not considered here, since we calculate diversity).

3. You need to calculate *mean alpha diversity* and *gamma diversity* within each locality (i.e. within each component of `gentry197` list). *Alpha diversity* of a transect is the number of species in that transect, and the *mean alpha diversity* is the mean number of species in ten transects per locality; *gamma diversity* is the total number of species in the dataset from a particular locality (i.e. number of species occurring at least once at the locality).

4. Draw the relationship of *mean alpha* and *gamma* diversity along latitude into one plot, add regression lines (linear regression), and also the labels (alpha and gamma). Note that

2

y-axis of the figure should be log-transformed.

5. Try to choose colours the most similar to the original figure in *Science* - you may choose the colour palette from http://colorbrewer2.org/.

6. Create the script in the way that it will directly load the data from this website, and save the figure as a **pdf file** (default size is ok). The pdf file should be named "gentry_XXXXXX.pdf", where XXXXXX is your student's ID.

Our final figure looks like this; you may try to get similar:
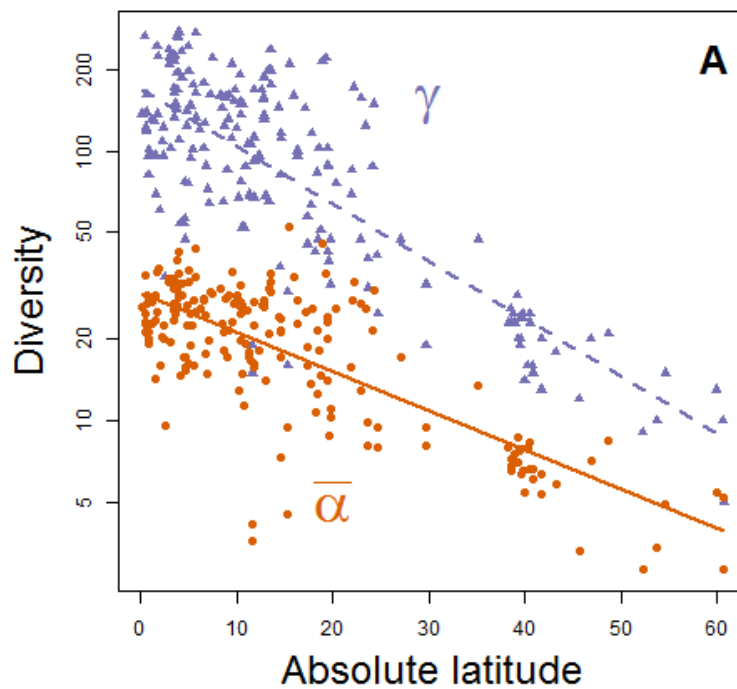

Figure 3: Final version of the pattern.

Hints

1. You need to upload two datasets - one using the function `load`, loading the list `gentry197`, and the other using `read.delim` loading the `gentry.coord` data frame with header (environmental) variables (you will need latitude from this data frame). Check https://www.davidzeleny.net/wiki/doku.php/recol:data_import#import_rdata_file to see how to download `*.RData` file directly from the Internet link.

2. You may check the structure of list by `str`, and the number of elements by `length`, names of the localities (elements) by `names`.

3. The easiest is to use the function `lapply` on the list - applying the function (defined in argument FUN) on each component of the list separately, and returning the list of results. You may also need functions like `colSums` or `rowSums`. Just note if you want to know how many species are there in one transect (row of data frame), you cannot simply sum the numbers in each row - that would return the number of individuals (of all species) in the transect, not the number of species. There is also prepared function `specnumber` in `library (vegan)`, which can be applied on the matrix of samples x species and returns the species richness (number of species) in each row (sample). To turn the resulting `list` object into `vector` of values, use `unlist`.

4. Here are some main problems:

    a. Log-transformation of the y-axis makes it a bit difficult with adding the regression line. You can achieve the log-transformation of y-axis by adding the argument `log = "y"` into `plot` function, to get axis labels on the y-axis in the original diversity values (numbers of species). When calculating the regression, calculate it on log-transformed diversity values (alpha or gamma), and later transform the predicted values using function `exp` back to the original not-log-transformed scale (because `exp (log (x)) == x`).

    b. Log-transformation of y-axis makes also more complicated to set up the range of y-axis values using `ylim` argument in `plot` function. This cannot start from zero, since logarithm of zero is minus infinity; you need to start from non-zero positive value. Check what is the lowest displayed diversity value (either gamma or mean alpha) and use this as the minimum value on the y-axis; similarly, you may check what is the highest value and use it as the maximum value.

    c. Adding the labels into the diagram - you need to add gamma (γ) and mean alpha ($\bar{\alpha}$). Use function `text` with coordinates of the position, and wrap the text into function `expression`. Check how to draw α letter with a bar above in `?plotmath`.

5. Regarding colours, the most similar seems to be the palette called *Dark2*. You need only two colours, but the http://colorbrewer2.org/ and RColorBrewer will force you to choose three - but you can still use just two from this selection.

6. Use `pdf` function and don't forget to use also `dev.off` in the end of the script.