

## dynGraph, the ins and outs

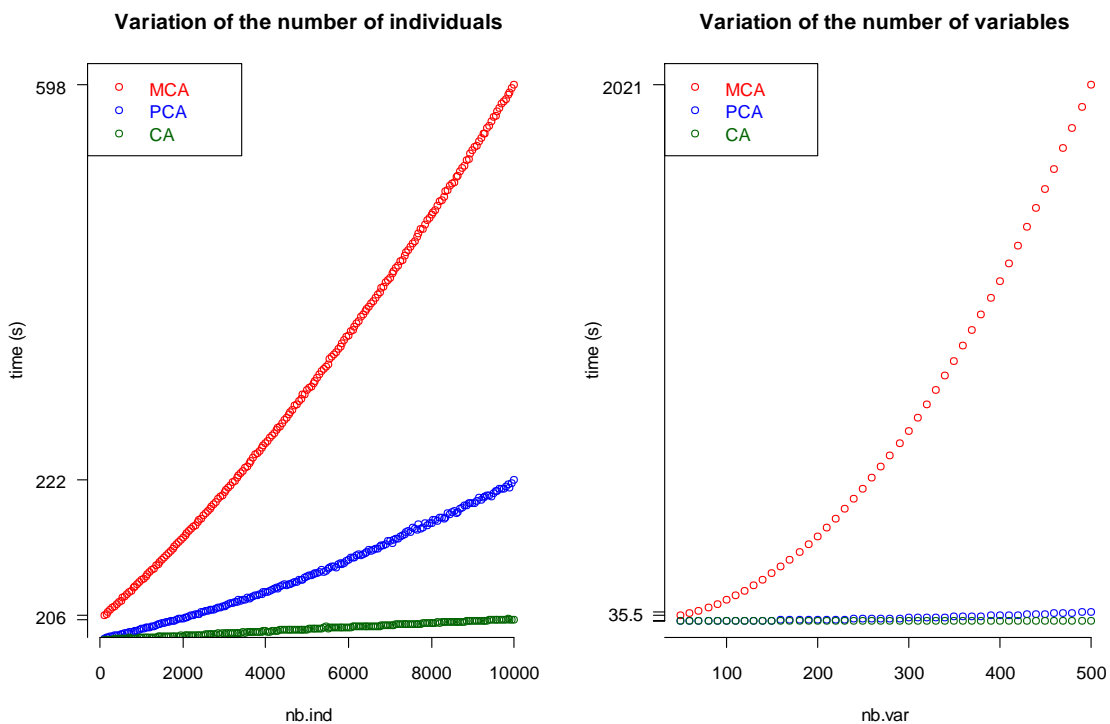
Sending information from R to Java meets its limits in the case of large data sets: the process is much longer and the final window is not generated.

*dynGraph*'s objective being a better visualization of data, we wanted to reduce the size of the object to plot, first to decrease the length of the process and second to stay within the limits where the java window is generated.

To study this limits, we performed several tests:

- Fix a small number of variables and make the number of individuals vary
- Fix a small number of individuals and make the number of variables vary

Here are the curbs of the process's length for CA, PCA and MCA when the numbers of individuals or variables vary:



Time and the number of individuals are almost linear whereas time grows with the squared number of variables.

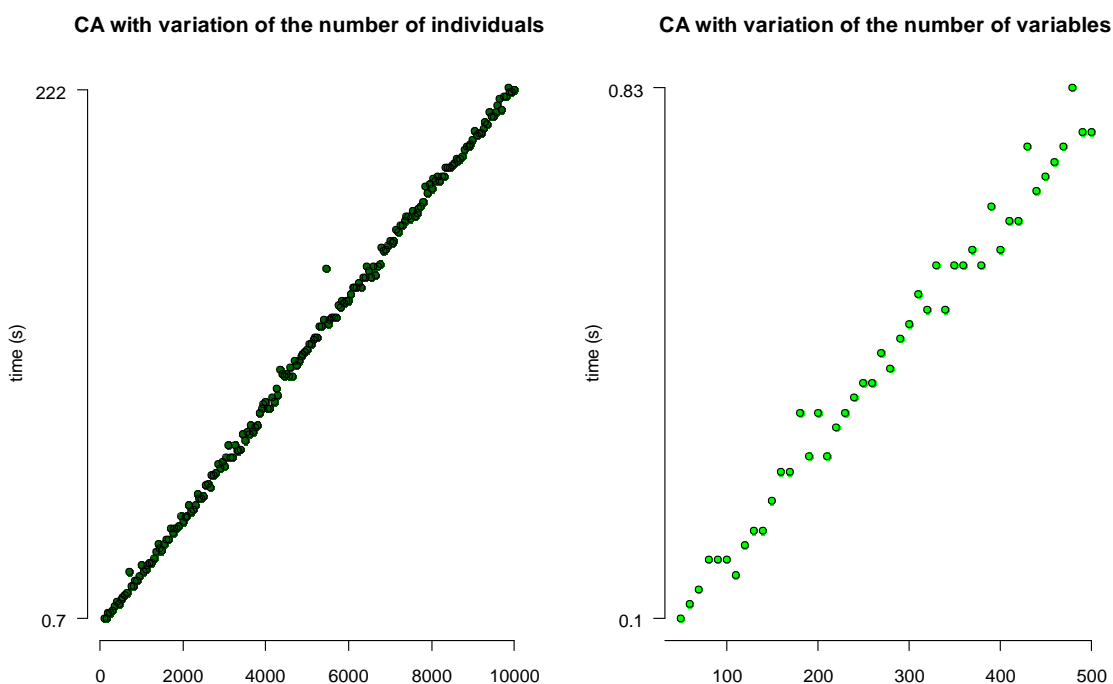
What we want is to fix the process's length to maximum 5mn. We created an internal function (*dynlis*) which uses the result of a CA, PCA or MCA and creates a new object of the same class but with restricted numbers of variables and individuals. For each method, we fixed:

- *nbrow* the maximum number of individuals (PCA, MCA) or rows (CA) to be kept. If there are some supplementary individuals or rows in the analysis, the reduced object has got a maximum of 20 supplementary individuals (or rows) and *nbrow*-20 active ones
- *nbobj* the maximum number of variables (PCA), categories (MCA) or columns (CA) to be kept. If there are supplementary variables of columns in the analysis, the final object has got a maximum of 20 supplementary variables (or columns or categories) and *nbobj*-20 active ones

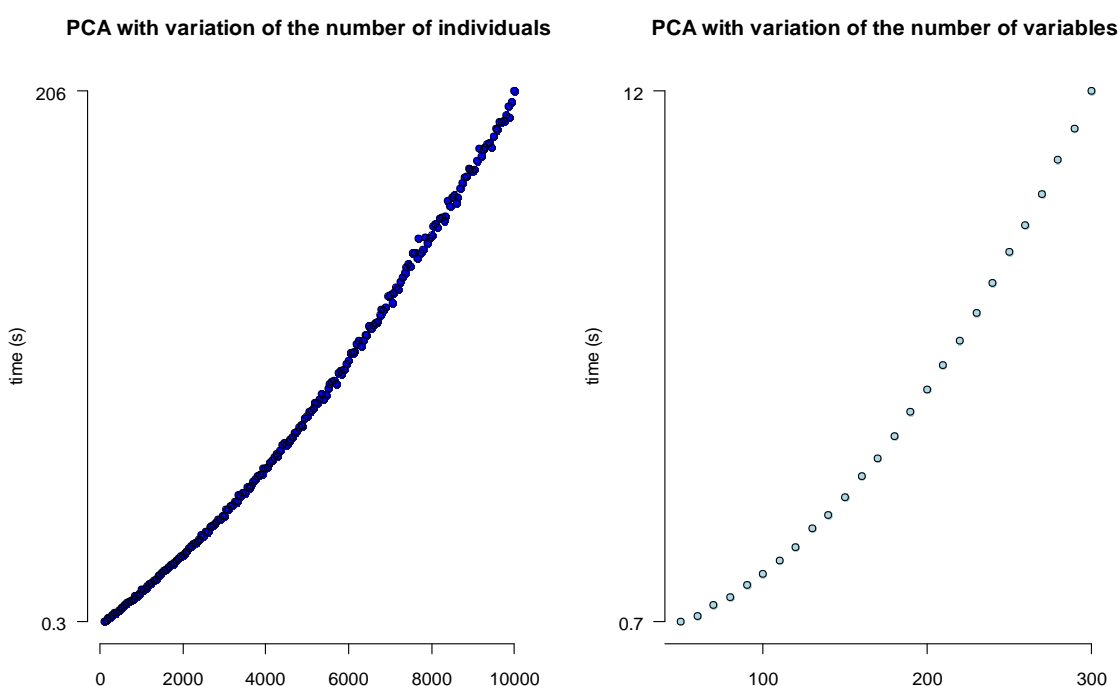
Individuals are selected according to their coordinates: individuals with high coordinates on at least one component of the analysis are selected. If the number of these selected individuals is less than *nbrow*, some other individuals are randomly added.

For CA and PCA, the function selects variables according to their coordinates. For MCA, categories are selected according to their v-test. Selection works in the same way as for individuals.

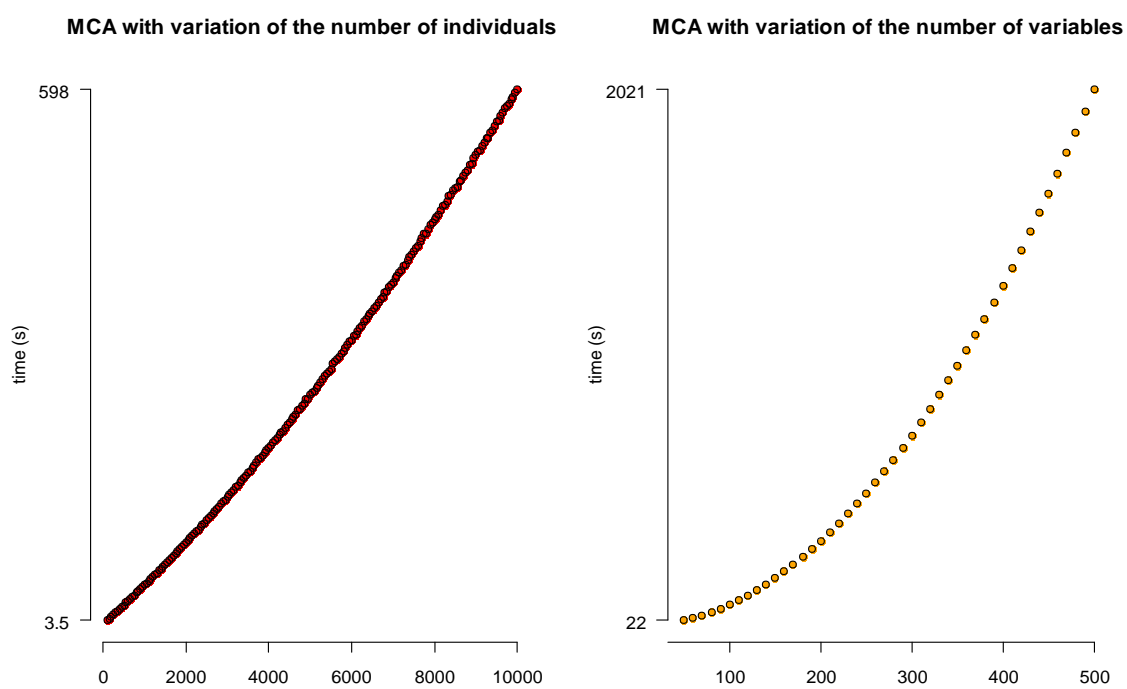
When performing *dynGraph*, the dimensions of your object are examined and, should they be over *nbrow* or *nbobj*, the new function, *dynlis*, is automatically called to create the restricted object.



For CA, the couple (*nbrow*; *nbobj*) has been fixed to (5000; 500) (or (500; 5000)).



For PCA, the maximum number of individuals has been fixed to 2000 and the maximum number of variables to 270.



For MCA, *nbrow* has been fixed to 200 and *nbobj* to 160.