How to use the EdgeClipper Algorithm

Jordan Rinder

Installing and Setting Up *BANJO*

1. Make sure that you have Java downloaded. If not, go to this website. You will need Java in order to run BANJO
2. Go to the BANJO website and request to download the software.
3. Once you have downloaded the BANJO software, it is be strongly advised to check that the software actually works.

```
* In order to do this, locate the *data* folder from the BANJO download folder
* Once in the data folder, go into the most recent release (release2.0 currently)
* Go to the static folder and locate the input folder. This folder will contain the files
necessary to test the BANJO software
* Go to the instruction on how to run BANJO from the website in which you downloaded the software,
and check to make sure everything works
```

Using Synthetic Data with *BANJO*

In R it is very easy to generate synthetic data. Based on the paper this software is a complement to, it is outline how to generate synthetic data

```r
# 9 variables with 500 observations
N <- 500
set.seed(12345)
A <- rnorm(N, mean = 0, sd = 1)
B <- rnorm(N, mean = 10, sd = 5)
C <- rnorm(N, mean = 0, sd = 10)
D <- A^3 + rnorm(N, mean = 0, sd = 0.1)
E <- ifelse(A + 10 >= B, A + rnorm(N, mean = 0, sd = 0.1), B/10 + rnorm(N, mean = 0, sd = 0.1))
EF <- (B - C) / (B + 10) + rnorm(N, mean = 0, sd = 0.1)
G <- ifelse(A + 10 >= B, A + sin(C) + rnorm(N, mean = 0, sd = 0.1),
            B  / 10 + sin(C) + rnorm(N, mean = 0, sd = 0.1))
H <- log(exp(A) + exp(EF)) + rnorm(N, mean = 0, sd = 0.1)
I <- (D + H) * (EF / 2) + rnorm(N, mean = 0, sd = 0.05)

x <- t(cbind(A, B, C, D, E, EF, G, H, I))


write(x, file = "input.data.txt", ncolumns = 9, sep = "    ")
```

This synthetic data generated can then be put into the format of the example files used when testing if BANJO works from step 3 part d. As an example, I write a text file separated by tabs (4 spaces) in order to make the synthetic data more visually appealing. It is the responsibility of the person working with BANJO to format the synthetic data properly so that BANJO can read in the file with no issues

Using the BANJO Output with the EdgeClipper Algorithm

For a great example of how to properly run and interpret the BANJO output, go here. For this algorithm, we will be looking at Static Bayesian Networks.

We take the BANJO output (making sure that it is in the proper directory), and feed it to several of the functions in funsEC.R. Through the help of these functions, we are able to calculate the B-values and C-values associated with the data.