Manual for fastCCLasso that estimates correlation matrix from compositional data in microbiome studies

Shen Zhang[1]

[1]School of Mathematical Sciences, Capital Normal University, Beijing 100048, China.

This is a short introduction to fastCCLasso based on a manuscript titled "fastCCLasso: Fast and efficient algorithm for estimating correlation matrix from compositional data." FastCCLasso can be used to explore the correlation network among microbes in microbiome studies.

The example OUT count data includes 58 samples and 61 OTUs (Srinivas et al. 2013). The example OUT file name is "example_input.csv" which uses "," as the separator of fields. A screenshot of the example data file is as follows.

```
"","OTU_1","OTU_2","OTU_3","OTU_4","OTU_5","OTU_6","OTU_7","OTU_8","OTU_9","OTU_10","OTU_11"
,"OTU_12","OTU_13","OTU_14","OTU_15","OTU_16","OTU_17","OTU_18","OTU_19","OTU_20","OTU_21","
OTU_22","OTU_23","OTU_24","OTU_25","OTU_26","OTU_27","OTU_28","OTU_29","OTU_30","OTU_31","OT
U_32","OTU_33","OTU_34","OTU_35","OTU_36","OTU_37","OTU_38","OTU_39","OTU_40","OTU_41","OTU_
42","OTU_43","OTU_44","OTU_45","OTU_46","OTU_47","OTU_48","OTU_49","OTU_50","OTU_51","OTU_52
","OTU_53","OTU_54","OTU_55","OTU_56","OTU_57","OTU_58","OTU_59","OTU_60","OTU_61"
"samp_1",81,0,15,459,42,3,5,45,32,3,10,10,0,2,119,4,4,40,82,4,13,6,32,20,31,30,206,0,21,43,8
,38,6,5,16,57,12,3,6,78,0,5,8,6,93,17,1,4,1,0,3,5,4,2,15,3,2,17,4,12,145
"samp_2",20,0,0,249,92,5,16,44,192,0,17,0,0,42,740,0,12,5,10,0,0,128,13,2,5,8,15,1,23,1,15,5
,0,1,1,379,106,3,1,3,2,3,34,13,76,25,0,4,0,2,0,3,4,17,4,2,0,6,2,1,23
```

Open an R terminal and input the following command

```
source("fastCCLasso_example.R");
```

The correlation coefficients for microbial interaction networks inferred by fastCCLasso, SparCC, CCLasso and COAT are output in the file "example_output.csv." A screenshot of the output data file is as follows. For user-defined input and output files, just replace file names in lines 9-10 in the file "fastCCLasso_example.R."

```
"variable_1","variable_2","fastCCLasso","SparCC","CCLasso","COAT"
"OTU_2","OTU_1",0,-0.143681398709718,0,-0.1028749345808
"OTU_3","OTU_1",0,0,0,0
"OTU_4","OTU_1",-0.0696445008505722,-0.372050572591351,-0.168759951991711,-0.369597967526781
"OTU_5","OTU_1",0,0.180397471322714,0,0.0821482271131679
"OTU_6","OTU_1",0.169344491755037,0.461875333666512,0.193280002319483,0.312704114096526
"OTU_7","OTU_1",0,-0.183009119335342,-0.12118144154586,-0.168058524380181
"OTU_8","OTU_1",0,0.214489929551278,0.135705828768634,0.110070143989227
"OTU_9","OTU_1",0.135138033702483,0.378684827454704,0.291308041742701,0.298569702112148
```

Reference

Srinivas, G., Möller, S., Wang, J., and al., et (2013), "Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering," *Nature Communications*, 4, 2462.

Zhang, S., Fang, H., and Hu, T. (2024), "fastCCLasso: Fast and efficient algorithm for estimating correlation matrix from compositional data," *Submitted*.