Module 5 Assignment

Daniel Bihnam

Load in and Process Data

```
load('C:/Downloads/tobacco_clr.Rdata')

data10 <- tobacco_clr$data[,1:10]
h10 <- tobacco_clr$H[,1:10]
color10 <- tobacco_clr$sample.color[1:10]
pch10 <- tobacco_clr$sample.pch[1:10]
otu10 <- tobacco_clr$otu.names[1:10]
```

First, I loaded in my data. Then, in order to specify the first 10 entries of the data, I used [1:10] and [,1:10] for normal vectors and matrix vectors.

Network with ppcor and igraph (Partial Correlation Analysis)

```
library(ppcor)

Warning: package 'ppcor' was built under R version 4.3.2

Loading required package: MASS

round(pcor(data10)$p.value,3)
```

	643311	4444760	687185	192795	164413	176468	255657	181589	219151	4425571
643311	0.000	0.294	0.012	0.621	0.395	0.002	0.837	0.414	0.661	0.006
4444760	0.294	0.000	0.170	0.205	0.096	0.001	0.772	0.706	0.010	0.324
687185	0.012	0.170	0.000	0.202	0.118	0.000	0.646	0.165	0.115	0.036
192795	0.621	0.205	0.202	0.000	0.762	0.904	0.865	0.668	0.032	0.793
164413	0.395	0.096	0.118	0.762	0.000	0.043	0.119	0.069	0.240	0.234
176468	0.002	0.001	0.000	0.904	0.043	0.000	0.783	0.085	0.610	0.002
255657	0.837	0.772	0.646	0.865	0.119	0.783	0.000	0.005	0.092	0.951
181589	0.414	0.706	0.165	0.668	0.069	0.085	0.005	0.000	0.261	0.524
219151	0.661	0.010	0.115	0.032	0.240	0.610	0.092	0.261	0.000	0.099
4425571	0.006	0.324	0.036	0.793	0.234	0.002	0.951	0.524	0.099	0.000

```
library(igraph)

Warning: package 'igraph' was built under R version 4.3.2


Attaching package: 'igraph'

The following objects are masked from 'package:stats':

    decompose, spectrum

The following object is masked from 'package:base':

    union

network <- (pcor(data10)$p.value<=0.05)*1

diag(network) <- 0
network.plot <- graph_from_adjacency_matrix(network)
#Create the plot
V(network.plot)$label <- otu10
#Use the first 10 OTUs to label the nodes
plot(network.plot,edge.arrow.size=.5,main='DB: 10 OTU Network 1')
```
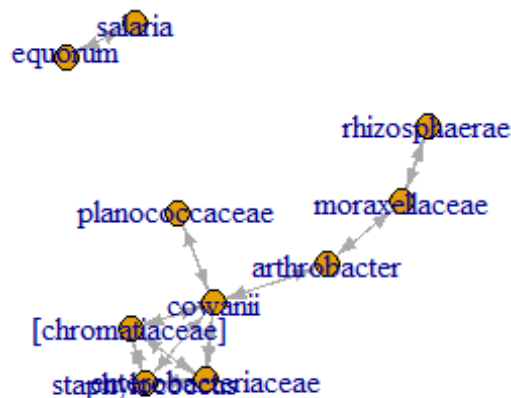

DB: 10 OTU Network 1

Using ppcor and igraph like we were shown in lecture, I created a network using a Gaussian graphical model. This model is able to display the direction of relationships between our first 10 OTUs in the dataset. We can see that there is a lot of relationship surrounding

cowanii, staphylococcus, chromatiacaea, and enterobacteriaceae. We can also see equorum and salaria are only related with each other, and not the other 8 OTUs. This network is the main network required for this assignment.

Network with glasso and qgraph (Graphical Lasso)
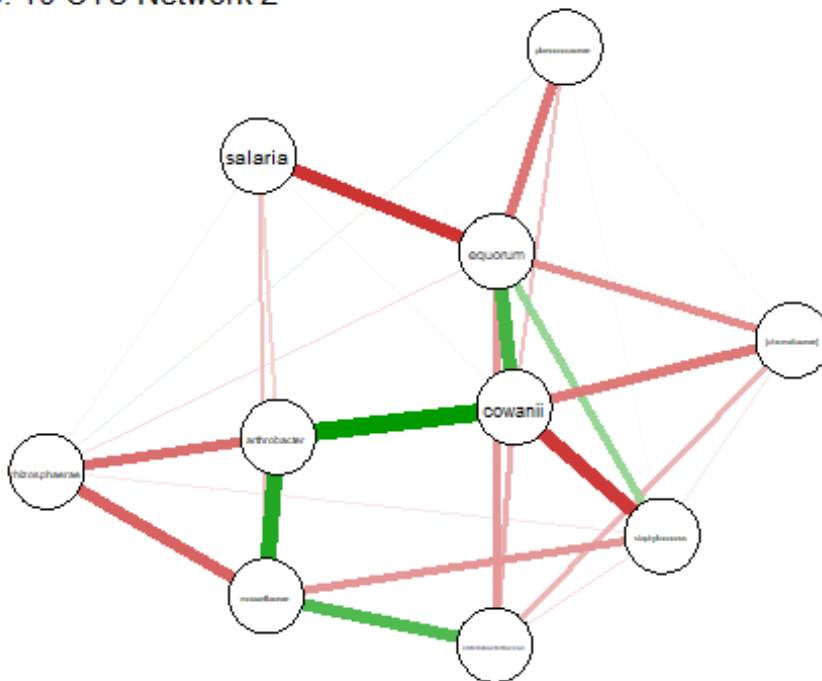
```
library(glasso)

cm <- cor(data10)
#Create correlation matrix
pcm <- glasso::glasso(cm,rho=0.1)
#Create partial correlation matrix using glasso library

library(qgraph)
```

Warning: package 'qgraph' was built under R version 4.3.2

```
gauss <- qgraph(pcm,layout='spring',labels=otu10,title='DB: 10 OTU Network 2')
```


DB: 10 OTU Network 2

```
#Create plot
```

While working on this assignment, I came across another package called glasso. Glasso reminded me of the lasso model that was used in our yellowdig discussion this week, so I decided to experiment with it and see how it differs from our previous model. It can be used to infer conditional dependencies between variables. This model tells a different story than the previous one, as we can see some bacteria such as equorum having a lot of

dependencies on and for other OTUs. Whereas in the first model, it was seperate from most of these OTUs. This was not required for this assignment, but I found it interesting and wanted to include it as an addition.