Module 5: Assignment

Aaron Hum

Stephanie Moyerman, PhD

LSC 541: Statistics for Biological Data Science I

2024-08-02

```
load("tobacco_clr.Rdata")

# subset the data to include the first 10 OTUs
data_subset <- tobacco_clr$data[, 1:10]
data_subset
```

```
     643311     4444760    687185     192795     164413     176468
RSL1 -1.1478018  4.4090263 -1.1478018 -1.1478018 -1.1478018  5.3772279
RSL3 -1.1994810  4.4317308 -1.1994810 -1.1994810 -1.1994810  5.6194430
RSL2 -1.1966517  4.0556217  3.5395467 -1.1966517 -1.1966517  5.3269106
RRD1 -2.3923418  3.4307041 -2.3923418 -2.3923418 -2.3923418  5.4147615
TPB3 -1.7247157  3.1194713 -1.7247157 -1.7247157 -1.7247157  3.1656334
NVB2 -1.3798340  4.3884869 -1.3798340 -1.3798340  4.4222843  5.3379706
DTL1 -1.8317339 -1.8317339 -1.8317339  3.5799122 -1.8317339  6.1895223
RRD2 -2.1666981  3.0150854  2.1373670 -2.1666981 -2.1666981  5.2746223
TPB2 -2.1883090  3.5924345 -2.1883090 -2.1883090 -2.1883090  4.4963027
RRD3  2.5201835  2.7484421 -2.5296725 -2.5296725 -2.5296725  5.2079438
NVB3 -1.3292235  3.7077291 -1.3292235 -1.3292235  4.6672286  5.7944493
DTL3 -1.5350387 -1.5350387 -1.5350387 -1.5350387 -1.5350387  5.8551427
DTL2 -1.7574065 -1.7574065 -1.7574065 -1.7574065 -1.7574065  6.2022188
TPB1 -2.1120912  3.5439006 -2.1120912 -2.1120912 -2.1120912  4.9645626
NVB1 -1.2249696  4.6912325 -1.2249696 -1.2249696  3.4289908  5.6986591
K6B2 -0.4585928 -0.4585928 -0.4585928 -0.4585928  4.2142361 -0.4585928
K6B1 -0.3850764 -0.3850764 -0.3850764 -0.3850764 -0.3850764 -0.3850764
L8B2 -0.4683688 -0.4683688 -0.4683688 -0.4683688  6.9389489 -0.4683688
L8B3 -0.4582567 -0.4582567 -0.4582567 -0.4582567  6.6005014 -0.4582567
TRS2 -0.6781239 -0.6781239 -0.6781239 -0.6781239 -0.6781239 -0.6781239
S4B1 -0.3564396 -0.3564396 -0.3564396 -0.3564396 -0.3564396 -0.3564396
R9B3 -0.4685848 -0.4685848 -0.4685848 -0.4685848  6.8683521 -0.4685848
TRS1 -0.6329046 -0.6329046 -0.6329046 -0.6329046 -0.6329046  4.3229225
TRS3 -0.5340613 -0.5340613 -0.5340613 -0.5340613 -0.5340613 -0.5340613
S4B3 -0.3228014 -0.3228014 -0.3228014 -0.3228014 -0.3228014 -0.3228014
C3B2 -0.3461644 -0.3461644 -0.3461644 -0.3461644 -0.3461644 -0.3461644
R9B1 -0.4067263 -0.4067263 -0.4067263 -0.4067263  6.5899552 -0.4067263
TB13 -1.1456869 -1.1456869  4.3096342 -1.1456869 -1.1456869 -1.1456869
K6B3 -0.4322439 -0.4322439 -0.4322439 -0.4322439  3.3519457 -0.4322439
S7B2 -0.3415675 -0.3415675 -0.3415675 -0.3415675 -0.3415675 -0.3415675
```

```
C3B3 -0.2672858 -0.2672858 -0.2672858 -0.2672858 -0.2672858 -0.2672858
R9B2 -0.4176403 -0.4176403 -0.4176403 -0.4176403  6.5744562 -0.4176403
TB22 -1.0574762 -1.0574762  3.6787223 -1.0574762 -1.0574762 -1.0574762
TB23 -1.0876274 -1.0876274  3.9230079 -1.0876274  3.4556674 -1.0876274
T5B2 -0.4862838 -0.4862838 -0.4862838 -0.4862838  6.8655160 -0.4862838
T5B1 -0.5194111 -0.5194111 -0.5194111 -0.5194111  7.1486826 -0.5194111
T5B3 -0.5091447 -0.5091447 -0.5091447 -0.5091447  6.5817651 -0.5091447
TB21 -0.9675910 -0.9675910  2.9836527 -0.9675910 -0.9675910 -0.9675910
L8B1 -0.3614232 -0.3614232 -0.3614232 -0.3614232  5.4131284 -0.3614232
TB11 -1.0916114 -1.0916114  3.4830996 -1.0916114 -1.0916114 -1.0916114
TB12 -0.8011263 -0.8011263  3.1108967 -0.8011263 -0.8011263 -0.8011263
S7B3 -0.3484924 -0.3484924 -0.3484924 -0.3484924 -0.3484924 -0.3484924
C3B1 -0.2899390 -0.2899390 -0.2899390 -0.2899390 -0.2899390 -0.2899390
S7B1 -0.2623762 -0.2623762 -0.2623762 -0.2623762 -0.2623762 -0.2623762
S4B2 -0.3394032 -0.3394032 -0.3394032 -0.3394032 -0.3394032 -0.3394032
     255657     181589     219151     4425571
RSL1 -1.1478018  5.8077908 -1.1478018 -1.1478018
RSL3 -1.1994810  6.0701357 -1.1994810 -1.1994810
RSL2 -1.1966517  5.5065364 -1.1966517 -1.1966517
RRD1 -2.3923418  4.2409767  2.7781422 -2.3923418
TPB3 -1.7247157  2.9757646  3.9277734  1.9388459
NVB2 -1.3798340  3.6953398  3.0628172 -1.3798340
DTL1 -1.8317339  5.8131855 -1.8317339 -1.8317339
RRD2 -2.1666981  4.5191628  2.5337823 -2.1666981
TPB2 -2.1883090  3.3088592  4.1467452  2.4751301
RRD3 -2.5296725  4.3760808  2.8030463 -2.5296725
NVB3 -1.3292235  4.0824226 -1.3292235 -1.3292235
DTL3 -1.5350387  5.4224587 -1.5350387 -1.5350387
DTL2 -1.7574065  5.8369747 -1.7574065 -1.7574065
TPB1 -2.1120912  3.5983358  4.6081289  2.9938542
NVB1 -1.2249696  4.1731931  3.7095044 -1.2249696
K6B2 -0.4585928  4.2863394 -0.4585928 -0.4585928
K6B1 -0.3850764  4.4824581 -0.3850764 -0.3850764
L8B2 -0.4683688 -0.4683688 -0.4683688 -0.4683688
L8B3 -0.4582567 -0.4582567 -0.4582567 -0.4582567
TRS2  4.7109478 -0.6781239 -0.6781239 -0.6781239
S4B1 -0.3564396 -0.3564396 -0.3564396 -0.3564396
R9B3 -0.4685848 -0.4685848 -0.4685848 -0.4685848
TRS1  4.5375794 -0.6329046 -0.6329046 -0.6329046
TRS3  4.7337969 -0.5340613 -0.5340613 -0.5340613
S4B3 -0.3228014 -0.3228014 -0.3228014 -0.3228014
C3B2 -0.3461644 -0.3461644 -0.3461644 -0.3461644
R9B1 -0.4067263 -0.4067263 -0.4067263 -0.4067263
TB13 -1.1456869  5.1587619 -1.1456869 -1.1456869
K6B3 -0.4322439  3.3289562 -0.4322439 -0.4322439
S7B2 -0.3415675 -0.3415675 -0.3415675 -0.3415675
C3B3 -0.2672858 -0.2672858 -0.2672858 -0.2672858
R9B2 -0.4176403 -0.4176403 -0.4176403 -0.4176403
TB22 -1.0574762  4.8800600 -1.0574762 -1.0574762
```

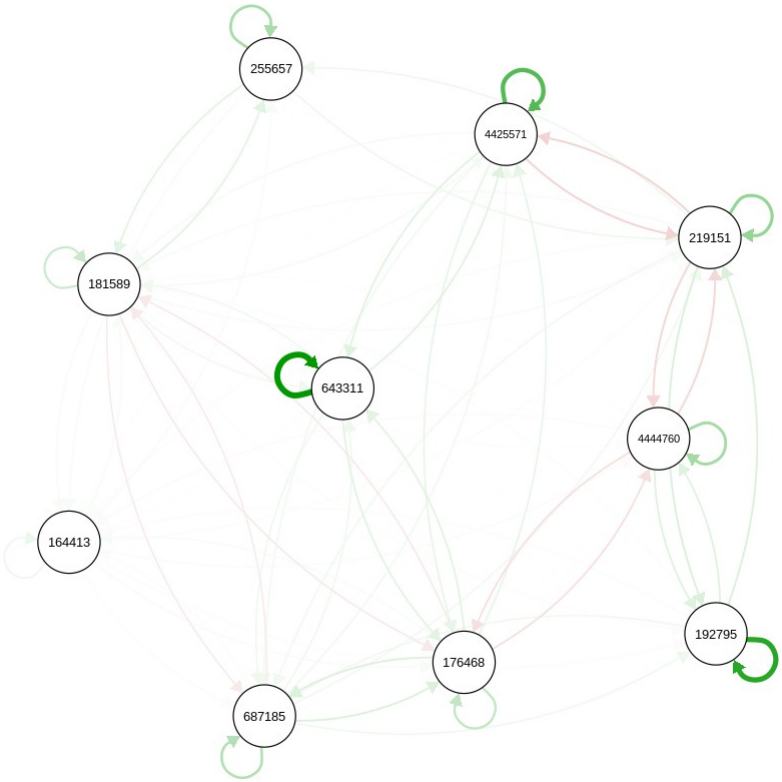
```
TB23 -1.0876274  5.0370560 -1.0876274 -1.0876274
T5B2 -0.4862838 -0.4862838 -0.4862838 -0.4862838
T5B1 -0.5194111 -0.5194111 -0.5194111 -0.5194111
T5B3 -0.5091447 -0.5091447 -0.5091447 -0.5091447
TB21 -0.9675910  4.4350864 -0.9675910 -0.9675910
L8B1 -0.3614232 -0.3614232 -0.3614232 -0.3614232
TB11 -1.0916114  4.5178604 -1.0916114 -1.0916114
TB12 -0.8011263 -0.8011263 -0.8011263 -0.8011263
S7B3 -0.3484924 -0.3484924 -0.3484924 -0.3484924
C3B1 -0.2899390  4.6445349 -0.2899390 -0.2899390
S7B1 -0.2623762 -0.2623762 -0.2623762 -0.2623762
S4B2 -0.3394032 -0.3394032 -0.3394032 -0.3394032

# load necessary packages
# install.packages("glasso")
library(glasso)
# install.packages("qgraph")
library(qgraph)

# calculate the covariance matrix
cov_matrix <- cov(data_subset)
cov_matrix

        643311      4444760    687185       192795       164413
176468
643311   0.63111786 -0.4845084 -0.060607931  0.123333259  0.63009539 -
1.0353239
4444760 -0.48450841  4.1201626 -1.099488832 -1.029576948 -1.13802002
3.9670235
687185  -0.06060793 -1.0994888  3.187831382  0.005397083 -0.37176143 -
2.2915335
192795   0.12333326 -1.0295769  0.005397083  0.860640701  0.68578581 -
0.8843171
164413   0.63009539 -1.1380200 -0.371761434  0.685785813 10.94770958 -
3.4510698
176468  -1.03532388  3.9670235 -2.291533489 -0.884317116 -3.45106979
8.1386891
255657   0.32140967 -1.1506849  0.037081388  0.391224984  0.42051766 -
1.6456569
181589  -0.89283599  2.0376490  0.563570325 -0.698532855 -3.79846100
4.3816342
219151  -0.28839638  2.2278574 -1.163651040 -0.835632050 -0.91516449
1.8627034
4425571 -0.12248916  0.1511638 -0.368565769 -0.052236320  0.08313653 -
0.6074064
        255657      181589     219151     4425571
643311   0.32140967 -0.8928360 -0.2883964 -0.12248916
4444760 -1.15068490  2.0376490  2.2278574  0.15116382
687185   0.03708139  0.5635703 -1.1636510 -0.36856577
192795   0.39122498 -0.6985329 -0.8356321 -0.05223632
```

```
164413   0.42051766 -3.7984610 -0.9151645  0.08313653
176468  -1.64565695  4.3816342  1.8627034 -0.60740639
255657   2.43093182 -2.1655117 -0.8387100  0.10174094
181589  -2.16551169  6.8549149  0.4177501 -0.80612232
219151  -0.83870998  0.4177501  2.8317758  0.68555385
4425571  0.10174094 -0.8061223  0.6855539  1.05979881
```

```
# fit the Gaussian graphical model using graphical lasso
glasso_fit <- glasso(cov_matrix, rho = 0.1)
glasso_fit
```

```
$w
             [,1]        [,2]        [,3]        [,4]       [,5]
[,6]
 [1,]  0.73111786 -0.46181735  0.03939238  0.12682034  0.5621217 -
0.9353229
 [2,] -0.46181735  4.22016257 -1.19949104 -0.92957157 -1.2380190
3.8670327
 [3,]  0.03939238 -1.19949104  3.28783138  0.10539706 -0.2717551 -
2.1915300
 [4,]  0.12682034 -0.92957157  0.10539706  0.96064070  0.5857832 -
0.8497197
 [5,]  0.56212168 -1.23801900 -0.27175508  0.58578317 11.0477096 -
3.3510890
 [6,] -0.93532293  3.86703274 -2.19153004 -0.84971972 -3.3510890
8.2386891
 [7,]  0.26531610 -1.05067877  0.11718924  0.31464344  0.5205122 -
1.5660149
 [8,] -0.79283806  1.93765864  0.46357282 -0.59853442 -3.6984700
4.2816294
 [9,] -0.23126306  2.12785357 -1.06365224 -0.73563118 -0.8151735
1.8576004
[10,] -0.02248898  0.08779529 -0.26856577 -0.07242801  0.1831340 -
0.5073892
             [,7]       [,8]       [,9]       [,10]
 [1,]  0.26531610 -0.7928381 -0.2312631 -0.02248898
 [2,] -1.05067877  1.9376586  2.1278536  0.08779529
 [3,]  0.11718924  0.4635728 -1.0636522 -0.26856577
 [4,]  0.31464344 -0.5985344 -0.7356312 -0.07242801
 [5,]  0.52051217 -3.6984700 -0.8151735  0.18313404
 [6,] -1.56601488  4.2816294  1.8576004 -0.50738917
 [7,]  2.53093182 -2.0655103 -0.7387166  0.09034248
 [8,] -2.06551026  6.9549149  0.5177495 -0.70612297
 [9,] -0.73871662  0.5177495  2.9317758  0.58555585
[10,]  0.09034248 -0.7061230  0.5855558  1.15979881

$wi
            [,1]        [,2]        [,3]        [,4]        [,5]
[,6]
 [1,] 1.70646426  0.00000000  0.10619439  0.00000000  0.00000000
```

```
0.18569604
 [2,] 0.00000000  0.60208630 -0.01112989  0.20617035 -0.03146500 -
0.20848204
 [3,] 0.10620161 -0.01115004  0.50643817  0.09938869  0.03107650
0.22757833
 [4,] 0.00000000  0.20620921  0.09938915  1.44405510 -0.02346325
0.00000000
 [5,] 0.00000000 -0.03146834  0.03107706 -0.02346322  0.12045880
0.04400977
 [6,] 0.18570671 -0.20851458  0.22757608  0.00000000  0.04400672
0.40384552
 [7,] 0.00000000  0.01529398  0.00000000  0.00000000  0.04356024
0.00000000
 [8,] 0.09297855 -0.01374790 -0.13104525  0.03009555  0.05460553 -
0.14557999
 [9,] 0.00000000 -0.25965514  0.09398462  0.23691995  0.03062859
0.00000000
[10,] 0.19553382  0.00000000  0.09379811  0.00000000  0.02273115
0.15317650
            [,7]        [,8]        [,9]       [,10]
 [1,] 0.00000000  0.09298378  0.00000000  0.19552876
 [2,] 0.01527024 -0.01376973 -0.25964480  0.00000000
 [3,] 0.00000000 -0.13104520  0.09398372  0.09379879
 [4,] 0.00000000  0.03009558  0.23692082  0.00000000
 [5,] 0.04356436  0.05460625  0.03062606  0.02273604
 [6,] 0.00000000 -0.14557885  0.00000000  0.15317156
 [7,] 0.57305413  0.18065049  0.11352134  0.00000000
 [8,] 0.18065040  0.34483449  0.05540745  0.06997103
 [9,] 0.11351761  0.05540500  0.70641800 -0.28038463
[10,] 0.00000000  0.06996858 -0.28038699  1.13531005

$loglik
[1] NA

$errflag
[1] 0

$approx
[1] FALSE

$del
[1] 9.554498e-05

$niter
[1] 3

# visualize the network
qgraph(glasso_fit$wi, labels = colnames(data_subset), layout =
"spring",
       vsize = 7, esize = 5)
```



Interpretation:

Strong Associations:

- Microbes identified by OTUs 643311, 192795, and 442571 have thicker green loops, indicating strong positive self-associations.
- There is a strong positive association between 643311 and 442571, indicated by the thick green edge connecting them.

Weak Associations:

- Several nodes are connected by thin green and red edges, indicating weaker positive and negative associations, respectively.

Negative Associations:

- Some red edges indicate negative associations, suggesting that as the abundance of one microbe increases, the abundance of the connected microbe tends to decrease.