Escherichia Coli Network

Example for GeneNet 1.2.7 (June 2013) or later

This note reproduces the "Escherichia coli" network example from J. Schäfer and K. Strimmer. 2005. *A shrinkage approach to large-scale covariance estimation and implications for functional genomics.* Statist. Appl. Genet. Mol. Biol. **4**: 32 (http://dx.doi.org/10.2202/1544-6115.1175)

Load GeneNet package

```
library("GeneNet")
```

```
## Loading required package: corpcor
## Loading required package: longitudinal
## Loading required package: fdrtool
## Loading required package: igraph
```

E. Coli data set (9 time points for 102 genes):

```
data(ecoli)
dim(ecoli)
```

```
## [1]   9 102
```

Estimation of partial correlations

Estimate matrix of partial correlation using a shrinkage estimator:

```
pc = ggm.estimate.pcor(ecoli)
```

```
## Estimating optimal shrinkage intensity lambda (correlation matrix): 0.1804
```

```
dim(pc)
```

```
## [1] 102 102
```

Assign p-values, q-values and empirical posterior probabilities to all 5151 potential edges in the network:

```
ecoli.edges = network.test.edges(pc, direct=TRUE, fdr=TRUE)
```

```
## Estimate (local) false discovery rates (partial correlations):
## Step 1... determine cutoff point
## Step 2... estimate parameters of null distribution and eta0
## Step 3... compute p-values and estimate empirical PDF/CDF
## Step 4... compute q-values and local fdr
## Step 5... prepare for plotting
```

```
##
## Estimate (local) false discovery rates (log ratio of spvars):
## Step 1... determine cutoff point
## Step 2... estimate parameters of null distribution and eta0
## Step 3... compute p-values and estimate empirical PDF/CDF
## Step 4... compute q-values and local fdr
## Step 5... prepare for plotting
```

```
dim(ecoli.edges)
```

```
## [1] 5151   10
```

The table lists all edges in the order strength of partial correlations:

```
ecoli.edges[1:5,]
```

```
##         pcor node1 node2         pval         qval      prob    log.spvar
## 1  0.2318566    51    53 2.220446e-16 3.612205e-13 1.0000000 -0.043537019
## 2  0.2240555    52    53 2.220446e-16 3.612205e-13 1.0000000 -0.040249854
## 3  0.2150782    51    52 2.220446e-16 3.612205e-13 1.0000000 -0.003287165
## 4  0.1732886     7    93 3.108624e-15 3.792816e-12 0.9999945 -0.025293430
## 5 -0.1341889    29    86 1.120813e-09 1.093998e-06 0.9999516  0.022305368
##    pval.dir  qval.dir prob.dir
## 1 0.3803869 0.7557272        0
## 2 0.4173922 0.7724561        0
## 3 0.9471949 0.8851073        0
## 4 0.6103234 0.8323249        0
## 5 0.6531371 0.8415749        0
```

Decide which edges to include in the network

To obtain a graph you need to select top ranking edges according to a suitable criterion. Here are some suggestions:

1. Use local fdr cutoff 0.2, i.e. include all edges with posterior probability of at least 0.8.

```
ecoli.net = extract.network(ecoli.edges)
```

```
##
## Significant edges:  125
##     Corresponding to  2.43 %  of possible edges
##
## Significant directions:  377
##     Corresponding to  7.32 %  of possible directions
## Significant directions in the network:  17
##     Corresponding to  13.6 %  of possible directions in the network
```

```
dim(ecoli.net)
```

```
## [1] 125  11
```

2. Use local fdr cutoff 0.1, i.e. i.e. include all edges with posterior probability of at least 0.9.

```
ecoli.net = extract.network(ecoli.edges, cutoff.ggm=0.9, cutoff.dir=0.9)
```

```
##
## Significant edges:  65
##     Corresponding to  1.26 %  of possible edges
##
## Significant directions:  269
##     Corresponding to  5.22 %  of possible directions
## Significant directions in the network:  6
##     Corresponding to  9.23 %  of possible directions in the network
```

```
dim(ecoli.net)
```

```
## [1] 65 11
```

3. Include a fixed number of edges, say the 70 strongest edges

```
ecoli.net = extract.network(ecoli.edges, method.ggm="number", cutoff.ggm=70)
```

```
##
## Significant edges:  70
##     Corresponding to  1.36 %  of possible edges
##
## Significant directions:  377
##     Corresponding to  7.32 %  of possible directions
## Significant directions in the network:  9
##     Corresponding to  12.86 %  of possible directions in the network
```

```
dim(ecoli.net)
```

```
## [1] 70 11
```

Plot network For plotting we use the igraph package (http://igraph.org). Note igraph is automatically installed with GeneNet.

```
library("igraph") #
```

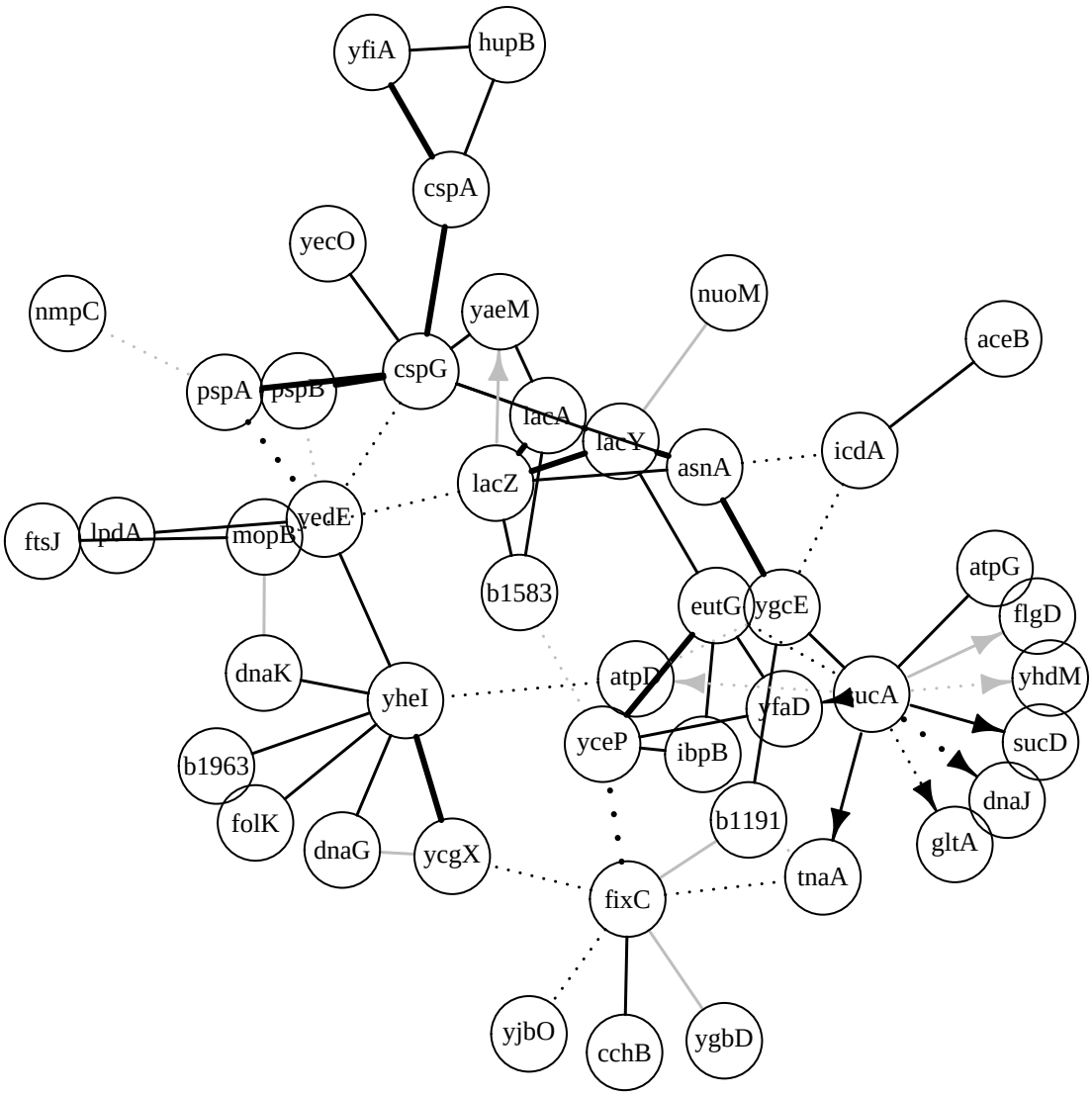
Create igraph object from the list of edges:

```
node.labels = colnames(ecoli)
igr1 = network.make.igraph(ecoli.net, node.labels)
igr1
```

```
## IGRAPH DN-- 46 70 --
## + attr: name (v/c), color (v/l), label.color (v/n), arrow.mode
##   (e/n), width (e/n), color (e/c), lty (e/n)
```

Plot the network:

```
plot(igr1, main="Ecoli Network", vertex.label.cex=0.8, edge.arrow.size=0.8)
```



It is also possible to produce a graph showing correlations as edge labels:

```
igr2 = network.make.igraph(ecoli.net, node.labels, show.edge.labels=TRUE)
```

```
plot(igr2, main="Ecoli Network with Partial Correlations as Edge Labels",
vertex.label.cex=0.8, edge.arrow.size=0.8)
```

Ecoli Network with Partial Correlations as Edge Labels