Graph clustering with the Stochastic block model (using Daudin's ICL)

Martin Metodiev

Contents


```
library(Matrix) # sparse matrix
library(FixedPoint) # Fixed point solution for the variational M-step
library(cluster) # Clustering to find the initial value
source("sbm_prototype.R")
source("classes.R")
set.seed(2134)
```

Daudin's Paper

The following is a brief project with the goal to implement Daudin's ICL:

$$\begin{aligned} \pi &\sim \text{Dirichlet}_K(\alpha), \\ \theta_{k,l} &\sim \text{Beta}(a_0, b_0), \\ Z_i &\sim \mathcal{M}(1, \pi), \\ \forall (i,j), \quad x_{ij} \mid Z_{ik} Z_{jl} = 1 &\sim \mathcal{B}(\theta_{k,l}). \end{aligned} \tag{1}$$

Simulation scenario

The following simulation is copied from a vignette from the greed package. However, none of its code is used.

We begin by simulating from a hierarchically structured SBM model, with 2 large clusters, each composed of 3 smaller clusters with higher connection probabilities, making a total of 6 clusters. Greed comes shipped with simulation function for the different generative models it handle and we will take advantage of the '`rsbm()` function to simulate an SBM graph with 6 clusters and 400 nodes:

```
N  <- 400           # Number of node
K  <- 6             # Number of cluster
pi <- rep(1/K,K)    # Clusters proportions
lambda   <- 0.1     # Building the connectivity matrix template
lambda_o <- 0.01
Ks <- 3
```

```
mu <- bdiag(lapply(1:(K/Ks), function(k){
  matrix(lambda_o,Ks,Ks)+diag(rep(lambda,Ks))}))+0.001
sbm <- sim_sbm(N,pi,mu) # Simulation
```

```
d_test = var_bayes_model_selection(sbm$x, Q=3)
#> [1] "ICL:  -7540.04840521306 for  3 clusters"
#> [1] "ICL:  -7433.89124634079 for  4 clusters"
#> [1] "ICL:  -7344.60699162115 for  5 clusters"
#> [1] "ICL:  -7308.1511059694 for  6 clusters"
#> [1] "ICL:  -7337.35087244629 for  7 clusters"
table(d_test@cl, sbm$cl)
#>
#>      1  2  3  4  5  6
#>   1  0  0  0  0  0 68
#>   2  0  0  0  0 60  0
#>   3  0  0  0 79  0  0
#>   4  1  2 64  0  0  0
#>   5  0 62  0  0  0  0
#>   6 62  2  0  0  0  0
```