

1. We use the following procedure for all three.

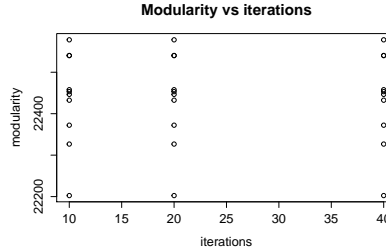
1. Find  $v$ , the first eigenvector for  $A - D, D^{-1/2}AD^{-1/2} - I, B$
2. Search over values  $c$ : compute the modularity of the partition  $\{i : v_i > c\} \cup \{i : v_i \leq c\}$ . Limit this search to values for  $c$  producing partitions with at least 20 nodes.
3. If the modularity of the best obtained partition exceeds the modularity of the entire graph, partition the graph and apply the procedure to each of the two subgraphs obtained. Otherwise, stop.

The results are as follows:

Matrix	No. clusters	Modularity
$A - D$	6	7415.6
$D^{-1/2}AD^{-1/2} - I$	47	1413.2
$B$	75	17205.1

The modularity matrix  $B$  gives the best result by far, while the normalized Laplacian is seen to be quite ill-suited for this problem.

2. All of the 10 initializations converged (reached a local minimum) within  $10n$  iterations.



The maximum modularity obtained was 22578.2, which exceeds the best obtained in step 1.

3.

We initialize the greedy algorithm with the clusters from step 1.

Matrix	No. clusters	Modularity
$A - D$	6	19038.8
$D^{-1/2}AD^{-1/2} - I$	47	20749.8
$B$	75	23328.1

In all cases the modularity increased substantially, and by using the cluster obtained from the modularity matrix  $B$  we can do better than the random initializations in part 2.