# Social Network Analysis

Exercises

1. List 3 of the different types of centrality we discussed in class.

2. What is the fundamental difference between eigenvector centrality and the other centrality measures? What does eigenvector centrality take into account that the other measures do not?

3. Compute the betweenness centrality of the shaded node in the following graph.
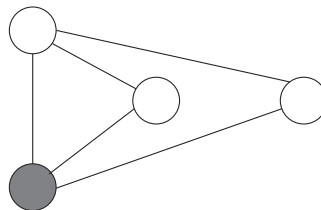
4. What is the purpose of normalizing measures of centrality?

5. How does the modularity maximization clustering algorithm define a cluster/community?

6. The term *spectral* clustering implies eigenvectors/eigenvalues are being used in some way. For what matrix are we calculating eigenvectors in these methods? The adjacency matrix? Do we use the eigenvectors associated with the largest eigenvalues, like in PCA?

7. Does the modularity maximization algorithm automatically determine the number of clusters? Do spectral methods that use k-means?

8. What clustering method that you've already learned about is the Minimum Spanning Tree method equivalent to?

9. For the following graph, highlight the minimum spanning tree. Is the minimum spanning tree unique? In other words, will everyone draw the same MST?
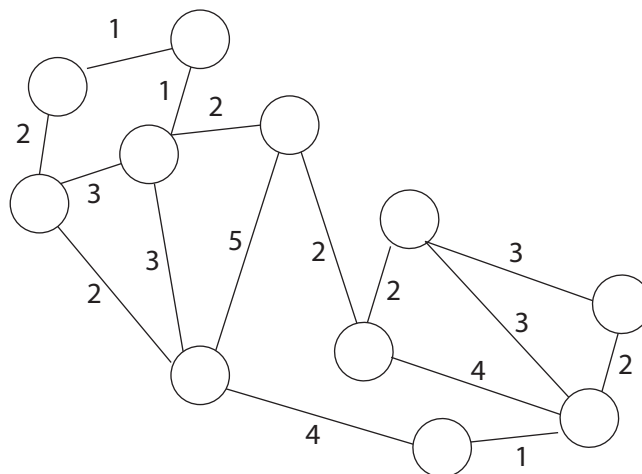
10. For the following hypothesis, indicate whether they are at the individual level, dyad level, or network level.

    a. Degree in an email network is related to tenure at a company.

    b. Density of a company's employee email network is related to stock price.

    c. Number of friends on facebook is a predictor of introversion or extroversion.

    d. The number of pages that two people both "like" is related to how long they've been friends.

    e. The amount of trade between two countries is related to the amount of communication between their highest governmental bodies.

    f. The military strength of a country is related to how many countries from whom they import goods.

11. Why do we turn to simulation to compute p-values for hypothesis tests on networks? Why can't we use traditional methods?

12. If we use the QAP approach to simulate a p-value for a dyadic level hypothesis, in essence what are we changing about the network in each iteration of that simulation? In other words, if I have a network and I perform a QAP permutation, what have I essentially done?

13. What does network autocorrelation entail? What types of attributes are involved?

14. Give an example of a question that could be answered by network autocorrelation.

List of Key Terms

Centralities

Network Centralization

Laplacian Matrix

Normalized Laplacian

Modularity

Spectral Clustering

Modularity Maximization

Minimum Spanning Trees

Simulating p-values

Individual Level Hypothesis

Dyadic Level Hypothesis

Network Autocorrelation

Network Regression