You may submit the homework in pairs. Write both full names when submitting.

1. Prove that a graph is connected if and only if its adjacency matrix W is irreducible.
2. Let G be an undirected and connected graph, with $W_{ii} = 0$ for all i.
   (a) Prove that $\lambda_n < 0$ (hint: consider $\mathrm{Tr}(W)$).
   (b) Assume that the graph is bipartite. Prove that the eigenvalues come in pairs such that $\lambda_1 = -\lambda_n, \lambda_2 = -\lambda_{n-1}$ etc.
3. Prove that the multiplicity of the 0 eigenvalue of the Laplacian is equal to the number of connected components in the graph.
4. Coding problem: Implement the stochastic block model we introduced in class for $n = 100$. Plot the first three eigenvalues and eigenvectors of L and compare to what we found in class for $\mathbb{E}[L]$ for
   - $p = 0.7$ and $q = 0$.
   - $p = 0.7$ and $q = 0.6$.
   - $p = 1$ and $q = 1$.
   - $p = 0$ and $q = 0.7$. (What kind of graph is this?)
5. Prove the following Proposition:

   Given a graph $G = (V;E;W)$ and a partition (S,S^c) of V , Ncut(S) corresponds to the probability, in the random walk associated with G, that a random walker in the stationary distribution goes to S^c conditioned on being in S plus the probability of going to S conditioned on being in S^c, more explicitly:

$$\mathrm{Ncut}(S) = \mathrm{Prob}\{X(t+1) \in S^c | X(t) \in S\} + \mathrm{Prob}\{X(t+1) \in S | X(t) \in S^c\}, \tag{1}$$

   where $\mathrm{Prob}\{X(t) = i\} = \frac{\deg(i)}{\mathrm{vol}(G)}$.
6. Prove the following proposition regarding the normalized cut on a graph with graph-Laplacian $L_G = D - W$ for two partitions S and S^c.

$$\mathrm{Ncut}(S) = y^T L_G y, \tag{2}$$

   where

$$y_i = \begin{cases} \left(\frac{\mathrm{vol}(S^c)}{\mathrm{vol}(S)\mathrm{vol}(G)}\right)^{1/2}, & i \in S \\ -\left(\frac{\mathrm{vol}(S)}{\mathrm{vol}(S^c)\mathrm{vol}(G)}\right)^{1/2}, & i \in S^c \end{cases} \tag{3}$$

7. Coding problem:
   - Implement spectral clustering and self-tuning spectral clustering.
   - Select two of the datasets in *Data6.mat* and apply k-means, spectral clustering and self-tuning spectral clustering [Zelnik-Manor &Perona 2005] to the datasets you selected.
   - Plot the results (2D scatter plots of the datasets with points colored by output labels) and report what parameters you used to construct the graphs.
   - Compare the performance of the different methods and explain why the methods perform well or poorly on the data.
8. Coding problem:

- Choose an image and reshape it to a modest size (few hundred pixels for width and length).
- Construct a graph using patches for features (extract patches as in HW1) or RGB values of the pixel or combination of both.
- Calculate the first 4 eigenvectors (corresponding to smallest eigenvalues) of the graph-Laplacian of your choice and plot them as reshaped images.
- Apply spectral clustering for 3 different k values and plot the results (labeled images).