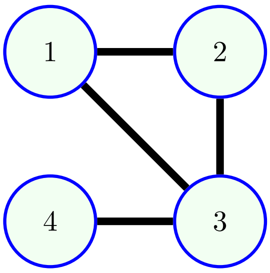
Network Science
Autumn 2022
Problem sheet 10



1. The modularity matrix for the graph shown above has the following eigenvalues and eigenvectors: $\lambda_1 = -1.57, \lambda_2 = -1, \lambda_3 = 0.319, \lambda_4 = 0$,

$$\mathbf{v}_1 = [-0.2054671, -0.2054671, 0.85011161, -0.43917741]^T$$

$$\mathbf{v}_2 = [7.07106781e-01, -7.07106781e-01, 6.89664944e-16, -3.72946812e-16]^T$$

$$\mathbf{v}_3 = [-0.4558325, -0.4558325, 0.16525815, 0.74640686]^T$$

$$\mathbf{v}_4 = [0.5, 0.5, 0.5, 0.5]^T$$

Explain how the spectral community detection algorithm will partition the graph into two communities.
Solution: The eigenvector corresponding to the largest eigenvalue (λ_3) is used to construct the partition. Negative components of $\mathbf{v}_3$ (nodes 1 and 2) are assigned to one group, while positive components (nodes 3 and 4) are assigned to the other.

2. Consider a complete N-node weighted graph with weight matrix, $\mathbf{W}$, and diagonal degree matrix $\hat{\mathbf{D}}$ where $\hat{D}_{ii} = \hat{k}_i = \sum_{j=1}^{N} W_{ij}$.

   (a) Show that the normalized Laplacian, $\hat{\mathbf{L}} = \hat{\mathbf{D}}^{-1/2}(\hat{\mathbf{D}} - \mathbf{W})\hat{\mathbf{D}}^{-1/2}$, has a zero eigenvalue and find the corresponding eigenvector
   Solution: Let $\tilde{\mathbf{L}} = \hat{\mathbf{D}} - \mathbf{W}$. Then, using the same reasoning as used previously for the usual Laplacian matrix, $\tilde{\mathbf{L}}\mathbf{z} = 0$ where $\mathbf{z}$ is a column vector of N ones. So, if $\hat{\mathbf{D}}^{-1/2}\mathbf{v} = \mathbf{z}$, then $\hat{\mathbf{L}}\mathbf{v} = 0$ and the normalized Laplacian has a zero eigenvalue with $v_j = \sqrt{\hat{k}_j}$ as the corresponding eigenvector.

   (b) Now consider a partition of the graph into two groups with $s_i = \pm 1$ indicating which group node i has been assigned to. Provide an expression for the normalized cut size in terms of $\mathbf{s}, \mathbf{z}, \mathbf{W}$, and $\hat{\mathbf{D}}$ where $\mathbf{z}$ is a column vector of N ones. The normalized cut size for a weighted graph is defined as: (wieghted cut size) $*\ (1/\hat{K}_a + 1/\hat{K}_b)$ where $\hat{K}_a$ and $\hat{K}_b$ are the total weighted degrees (computed from $\hat{k}_i$) of the two groups of nodes in the partition.
   Solution: The weighted cut size is, $\hat{c} = 1/4 \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}(1 - s_i s_j) = 1/4(\hat{K} - \mathbf{s}^T\mathbf{W}\mathbf{s})$ where $\hat{K} = \text{trace}\left(\hat{\mathbf{D}}\right)$. The degree of group a is $K_a = 1/2 \sum_{i=1}^{N} \hat{D}_{ii}(1 + s_i) = 1/2(\hat{K} + \mathbf{z}^T\hat{\mathbf{D}}\mathbf{s})$. Similarly, $K_b = 1/2(\hat{K} - \mathbf{z}^T\hat{\mathbf{D}}\mathbf{s})$. So, the normalized cut size is $1/2(\hat{K} - \mathbf{s}^T\mathbf{W}\mathbf{s})\left[1/(\hat{K} + \mathbf{z}^T\hat{\mathbf{D}}\mathbf{s}) + 1/(\hat{K} - \mathbf{z}^T\hat{\mathbf{D}}\mathbf{s})\right]$.

(c) It can be shown that minimizing the normalized cut size is equivalent to finding $\mathbf{y}$ such that $\frac{\mathbf{y}^T\hat{\mathbf{L}}\mathbf{y}}{\mathbf{y}^T\hat{\mathbf{D}}\mathbf{y}}$ is minimized with $y_i \in \{1, -b\}$ and $\mathbf{y}^T\hat{\mathbf{D}}\mathbf{z} = 0$. Here $\hat{\mathbf{L}} = \hat{\mathbf{D}} - \mathbf{W}$, $\mathbf{y} = (\mathbf{z}+\mathbf{s}) - b(\mathbf{z}-\mathbf{s})$, and $b = \hat{K}_a/\hat{K}_b$. Find a transformation for $\mathbf{y}$ that results in the following equivalent minimization problem: Find $\mathbf{x}$ such that $\frac{\mathbf{x}^T\tilde{\mathbf{L}}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$ is minimized. Typically the 1st constraint on $\mathbf{y}$ ($y_i \in \{1, -b\}$) is removed, leaving one constraint for both $\mathbf{x}$ and $\mathbf{y}$. What is this constraint for $\mathbf{x}$?

Solution: $\mathbf{x} = \hat{\mathbf{D}}^{1/2}\mathbf{y}$ transforms the problem as required, and the constraint, $\mathbf{y}^T\hat{\mathbf{D}}\mathbf{z} = 0$ becomes, $\mathbf{x}^T\hat{\mathbf{D}}^{1/2}\mathbf{z} = 0$

(d) Find the vector $\mathbf{x}$ that solves the minimization problem described in (c) with the constraint that you derived. Provide an interpretation of your result.

Solution: From (a), we know that $\hat{\mathbf{D}}^{1/2}\mathbf{z}$ is an eigenvector of $\hat{\mathbf{L}}$. Since $\hat{\mathbf{L}}$ is symmetric positive semidefinite, the solution is the eigenvector corresponding to the smallest positive eigenvalue of $\hat{\mathbf{L}}$ (note that this eigenvector will be orthogonal to $\hat{\mathbf{D}}^{1/2}\mathbf{z}$). To apply this result, we have to transform the vector using $\mathbf{y} = \hat{\mathbf{D}}^{-1/2}\mathbf{x}$ and then comparing the elements to 1 and $-b$ provides an indication of which group the corresponding node should be placed in. To actually construct the partition, an additional rule is required.

(e) The minimization problem in (c) is typically solved for image segmentation applications (e.g. problems related to artificial vision). The spectral clustering method is based on computing the eigenvectors corresponding to the *largest* eigenvalues of $\hat{\mathbf{A}} = \hat{\mathbf{D}}^{-1/2}\mathbf{W}\hat{\mathbf{D}}^{-1/2}$. How are the eigenvalues and eigenvectors of $\hat{\mathbf{A}}$ and $\hat{\mathbf{L}}$ related?

Solution: Since $\hat{\mathbf{L}} = \mathbf{I} - \hat{\mathbf{A}}$, if $\hat{\mathbf{A}}\mathbf{x} = \lambda\mathbf{x}$, then $\hat{\mathbf{L}}\mathbf{x} = \mathbf{I}\mathbf{x} - \hat{\mathbf{A}}\mathbf{x} = \mathbf{I}\mathbf{x} - \lambda\mathbf{x} = (\mathbf{z} - \lambda)\mathbf{x}$, so if $\mathbf{x}$ is an eigenvector of $\hat{\mathbf{A}}$ with eigenvalue, λ, then it is also an eigenvector of $\hat{\mathbf{L}}$ with eigenvalue $1 - \lambda$.

3. Consider an undirected connected weighted graph with weight matrix $\mathbf{W}$ which has all elements non-negative.

(a) Propose a definition for a step of a random walk on this graph in terms of $\mathbf{W}$ and quantities derived from $\mathbf{W}$

Solution: A sensible approach is to connect T_{ij}, the probability of a step from i to j to the weight W_{ij}. Here, I will define $T_{ij} = W_{ij}/\hat{k}_i$ with $\hat{k}_i = \sum_{j=1}^{N} W_{ij}$.

(b) What is the stationary distribution for your model?

Solution: The derivation is essentially the same as that for RWGs (lecture 11): $p_{\infty,i} = \hat{k}_i/\hat{K}$ with $\hat{K} = \sum_{i=1}^{N} \hat{k}_i$.

(c) Describe (give a sketch rather than a rigorous argument) how the locations of walkers at large times are related to non-trivial graph partitions obtained from the normalized minimum cut problem discussed in 2(c) (after the relaxation of the first constraint on $\mathbf{y}$).

Solution: Let $\mathbf{p}^{(l)}$ be the probability vector for the random walk at step l. Then, as with RWGs, we set $\mathbf{w}^{(l)} = \mathbf{p}^{(l)}\hat{\mathbf{D}}^{-1/2}$, and $\mathbf{y}^{(l)} = \mathbf{w}^{(l)}\mathbf{V}$ where $\mathbf{V}$ is the eigenvector matrix for, $\hat{\mathbf{A}} = \hat{\mathbf{D}}^{-1/2}\mathbf{W}\hat{\mathbf{D}}^{-1/2}$. At very long times when the the distribution has not yet reached its stationary state, the solution for $\mathbf{w}$ will be a weighted sum of the eigenvectors corresponding to $\lambda_1 = 1$ and λ_2 (with $\lambda_2 < \lambda_1$). The eigenvector corresponding to λ_2 ($\mathbf{v}_2$) can be interpreted as the departure from the stationary state (or a perturbation to the stationary state). Since eigenvectors of $\hat{\mathbf{A}}$ are also eigenvectors of $\hat{\mathbf{L}}$ (shown in 2 (e)), and since subtracting λ_2 from one gives the smallest positive eigenvalue of $\hat{\mathbf{L}}$, the departure from the stationary state is also the eigenvector used to construct the partition using the image segmentation method described in question 2.