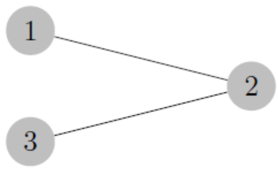
Contributors: burak.kurucu@uni-potsdam.de; mikhail.stepanov@uni-potsdam.de; mohammad.yakub@uni-potsdam.de; md.shafayet.hossen.chowdhury@uni-potsdam.de

*10. Problem sheet for* ***Statistical Data Analysis***

*Exercise 1*




*L(G): Laplacian Matrix*

*D(G): Degree Matrix*

*W(G): Affinity Matrix which is A(G): Adjacency Matrix in this exercise since* $\omega_{ij} = 1$ *for all ij.*

$$D = \begin{pmatrix} d(1) & 0 & 0 \\ 0 & d(2) & 0 \\ 0 & 0 & d(3) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$L = D - A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

$$L \cdot v = \lambda \cdot v \Rightarrow L \cdot v - \lambda \cdot v = 0 \Rightarrow (L - \lambda \cdot I) \cdot v = 0$$

*If v is nonzero, equation only have a solution if* $|L - \lambda \cdot I| = 0$

$$L - \lambda \cdot I = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{pmatrix}$$

$$det(L - \lambda \cdot I) = \begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = \left| \begin{array}{ccc|cc} 1-\lambda & -1 & 0 & 1-\lambda & -1 \\ -1 & 2-\lambda & -1 & -1 & 2-\lambda \\ 0 & -1 & 1-\lambda & 0 & -1 \end{array} \right.$$

$$= (1-\lambda)(2-\lambda)(1-\lambda) + 0 + 0 - (1-\lambda) - (1-\lambda)$$

$$= (-\lambda^3 + 4\lambda^2 - 5\lambda + 2) - (2 - 2\lambda) = -\lambda^3 + 4\lambda^2 - 3\lambda = -\lambda(\lambda^2 - 4\lambda + 3)$$

$$= -\lambda(\lambda - 3)(\lambda - 1) = 0$$

$$\Rightarrow \lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 3$$

$$(L - \lambda \cdot I) \cdot v = 0$$

$$for\ \lambda_1 = 0$$

$$(L - \lambda_1 \cdot I) \cdot v = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$x_1 = x_2, \qquad x_2 - x_3, \qquad \begin{cases} x_1 - x_2 = 0 \\ -x_1 + 2x_2 - x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}$$

$$\Longrightarrow X = \left( x_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) \Longrightarrow v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$for\ \lambda_2 = 1$$

$$(L - \lambda_2 \cdot I) \cdot v = \begin{pmatrix} 1-1 & -1 & 0 \\ -1 & 2-1 & -1 \\ 0 & -1 & 1-1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$-x_1 = x_3, \qquad x_2 = 0, \qquad \begin{cases} -x_2 = 0 \\ -x_1 + x_2 - x_3 = 0 \\ -x_2 = 0 \end{cases}$$

$$\Longrightarrow X = \left( x_3 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right) \Longrightarrow v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$for\ \lambda_3 = 3$$

$$(L - \lambda_3 \cdot I) \cdot v = \begin{pmatrix} 1-3 & -1 & 0 \\ -1 & 2-3 & -1 \\ 0 & -1 & 1-3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$x_1 = -\frac{1}{2} x_2, \qquad x_3 = -\frac{1}{2} x_2, \qquad x_1 = x_3, \qquad \begin{cases} -2x_1 - x_2 = 0 \\ -x_1 - x_2 - x_3 = 0 \\ -x_2 - 2x_3 = 0 \end{cases}$$

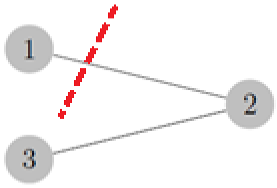
$$\Longrightarrow X = \left( x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right) \Longrightarrow v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

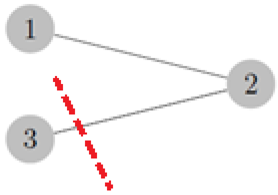
$$RatioCut(A_1, A_2, \ldots, A_k) = \sum_{i=1}^{k} \frac{cut(A_i, \overline{A_i})}{|A_i|}$$

$$nCut(A_1, A_2, \ldots, A_k) = \sum_{i=1}^{k} \frac{cut(A_i, \overline{A_i})}{vol(A_i)}$$

*Case 1: 2 clusters:*

*(i)* $A_1 = \{1\}, A_2 = \{2,3\}$ *(ii)* $A_1 = \{1,2\}, A_2 = \{3\}$




*(i)*

$$RatioCut(A_1, A_2) = \sum_{i=1}^{k} \frac{cut(A_i,\overline{A_i})}{|A_i|} = \frac{cut(A_1,\overline{A_1})}{|A_1|} + \frac{cut(A_2,\overline{A_2})}{|A_2|} = \frac{1}{1} + \frac{1}{2} = \frac{3}{2}$$

$$nCut(A_1, A_2) = \sum_{i=1}^{k} \frac{cut(A_i,\overline{A_i})}{vol(A_i)} = \frac{cut(A_1,\overline{A_1})}{vol(A_1)} + \frac{cut(A_2,\overline{A_2})}{vol(A_2)} = \frac{1}{1} + \frac{1}{3} = \frac{4}{3}$$

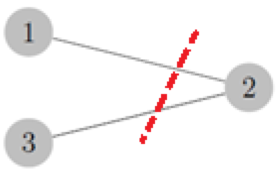
*(ii)*

$$RatioCut(A_1, A_2) = \sum_{i=1}^{k} \frac{cut(A_i,\overline{A_i})}{|A_i|} = \frac{cut(A_1,\overline{A_1})}{|A_1|} + \frac{cut(A_2,\overline{A_2})}{|A_2|} = \frac{1}{2} + \frac{1}{1} = \frac{3}{2}$$

$$nCut(A_1, A_2) = \sum_{i=1}^{k} \frac{cut(A_i,\overline{A_i})}{vol(A_i)} = \frac{cut(A_1,\overline{A_1})}{vol(A_1)} + \frac{cut(A_2,\overline{A_2})}{vol(A_2)} = \frac{1}{3} + \frac{1}{1} = \frac{4}{3}$$

*Case 2: 3 clusters:*

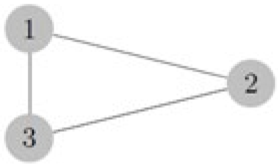
$A_1 = \{1\}, A_2 = \{2\}, A_3 = \{3\}$




$$RatioCut(A_1, A_2) = \sum_{i=1}^{k} \frac{cut(A_i,\overline{A_i})}{|A_i|} = \frac{cut(A_1,\overline{A_1})}{|A_1|} + \frac{cut(A_2,\overline{A_2})}{|A_2|} + \frac{cut(A_3,\overline{A_3})}{|A_3|} = \frac{1}{1} + \frac{2}{1} + \frac{1}{1} = 4$$

$$nCut(A_1, A_2) = \sum_{i=1}^{k} \frac{cut(A_i,\overline{A_i})}{vol(A_i)} = \frac{cut(A_1,\overline{A_1})}{vol(A_1)} + \frac{cut(A_2,\overline{A_2})}{vol(A_2)} + \frac{cut(A_3,\overline{A_3})}{vol(A_3)} = \frac{1}{1} + \frac{2}{2} + \frac{1}{1} = 3$$

*Furthermore: If we add a path from node 1 to 3 we will have strongly connected graph.*




*In that case:*

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$L = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

*Then* $\lambda_1 = 0, \lambda_2 = 3, \lambda_3 = 3$. *Still there is an eigenvalue equal to 0 but other eigenvalues are equivalent and 3. The "zero" eigenvalue tells us whether the graph is connected.*