Exercise 2: Implementing k-means and spectral clustering

(implementation exercise)

Exercise 3: Design your own exam questions

1. Please explain the difference of kernel-PCA(Principal Component Analysis) and PCA. What is the advantage to use kernel?

2. Please judge True or False for following each sentences.
[ ]Principal component analysis does not change the value of the main component depending on whether or not to standardize the data.
[ ]The principal component vector is the eigenvector of magnitude 1 of the correlation matrix
[ ]Principal component analysis is a statistical process that compresses multidimensional information into low dimensional information

3. Please divide into two cluster by k-means.

Name	maximal blood pressure	cigarette/Day
A	80	5
B	60	3
C	160	8
D	140	6
E	100	6
F	200	10

Exercise 1: Spectral graph theory

Because the Latex Layout function does not correct work, I show the allocation as following.
a) Picture 1 - 4
b) Picture 5 - 10
c) Picture 11 - 14
d) Picture 15 - 17

(a)

(b)

(c)

(d)

$$D = \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & & \ddots & 0 \\ 0 & & & d_n \\ 0 & 0 & \cdots & \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \\ a_{n1} & & a_{nn} \end{bmatrix}$$

$$L = D^{-1/2}(D-A)D^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{d_1}} & & & \\ & \frac{1}{\sqrt{d_2}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{d_n}} \end{bmatrix} \begin{bmatrix} \\ (D-A) \\ \\ \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{d_1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{d_n}} \end{bmatrix}$$

Gershgorin circle

$$L = \begin{array}{c|c} & \frac{1}{\sqrt{d_1}} \quad \cdots \quad \frac{1}{\sqrt{d_n}} \\ \hline \frac{1}{\sqrt{d_1}} & \begin{bmatrix} d_1 - a_{11} & -d_{12} & \cdots & -a_{1n} \\ \vdots & \ddots & & \\ a_{n1} & & & d_n - a_{nn} \end{bmatrix} \\ \vdots & \\ \frac{1}{\sqrt{d_n}} & \end{array}$$

$$D\left(\frac{d_1 - a_{11}}{d_1}, R_1\right) = D\left(\frac{d_1 - a_{11}}{d_1}, \sum |a_{1j}|\right) = D\left(\frac{d_1 - a_{11}}{d_1}\right.$$

$$D\left(\frac{d_i}{d_i}, \underbrace{\frac{1}{d_\alpha} + \frac{1}{d_\beta} + \cdots + \frac{1}{d_r}}_{\text{each degree is the degree of the vertex to which it is connected to}}\right)$$

$$\Downarrow$$

$$1$$

According to this method; given $D(D^{(1)}, D^{(2)})$; the EV are bounded by $D^{(1)} \pm D^{(2)}$

since the lowest value is 0; $1 - D^{(2)} = 0$; $D^{(2)} = 1$

the possible highest value is 2; $D^{(1)} + D^{(2)} = 1 + 1 = 2$ //

$$D = \begin{bmatrix} n-1 & 0 & 0 & 0 & 0 \\ 0 & n-1 & & 0 & 0 \\ 0 & 0 & \ddots & & \\ 0 & 0 & 0 & \ddots & \\ 0 & 0 & & & n-1 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & \ddots & & & 1 \\ 1 & & & & & \\ 1 & 1 & & & & 1 \\ 1 & & & & & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} \frac{1}{\sqrt{n-1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{n-1}} \end{bmatrix} \begin{bmatrix} n-1 & -1 & -1 & \cdots & -1 \\ -1 & \ddots & & & \\ \vdots & & \ddots & & \\ & & & \ddots & \\ -1 & & -1 & & n-1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{n-1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{n-1}} \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & \frac{-1}{n-1} & \frac{-1}{n-1} & \cdots & \frac{-1}{n-1} \\ -1 & 1 & \ddots & & \vdots \\ \vdots & & & & \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & \frac{-1}{n-1} & \frac{-1}{n-1} & \cdots & \frac{-1}{n-1} \\ \frac{-1}{n-1} & 1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \frac{-1}{n-1} & \cdots & \cdots & & 1 \end{bmatrix}$$

$$L = \begin{vmatrix} n-1 & -1 & -1 & \cdots & -1 \\ -1 & n-1 & & & \\ -1 & & n-1 & & \\ -1 & & & \ddots & \\ -1 & & & & \\ -1 & -1 & -1 & -1 & n-1 \\ \vdots & & & & \end{vmatrix}$$

According to proposition 22 the multiplicity of the EV 0 is the # of all connected components; since is all connected there is 1 component

b)

$$L - 0I = \left(\frac{1}{n-1}\right) \begin{vmatrix} n-1 & -1 & -1 & -1 & \cdots \\ -1 & n-1 & & & \\ -1 & & n-1 & & \\ -1 & & & \ddots & \\ \vdots & & & & \end{vmatrix}$$

$\downarrow$

$\det(L) = 0$

since the sum of all elements in its ~~row~~ is 0

$\lambda = 0$ is singular; that means

... connected components ; since is all connected there is 1 component

$\begin{pmatrix} -1 & -1 & -1 & -1 & n-1 \end{pmatrix}$

$$\mathbb{L} - 0I = \left(\frac{1}{n-1}\right) \begin{pmatrix} n-1 & -1 & -1 & -1 & \cdots \\ -1 & n-1 & & & \\ -1 & & n-1 & & \\ -1 & & & \ddots & \\ \vdots & & & & \ddots \\ & & & & n-1 \end{pmatrix}$$

$\downarrow$

$\det(\mathbb{L}) = 0$

since the sum of all elements in its row is 0

$\lambda = 0$ is singular ; that means has multiplicity 1

$$\mathbb{L} - nI = \frac{1}{n-1} \begin{pmatrix} -1 & -1 & -1 & \cdots & -1 \\ -1 & & & & \vdots \\ \vdots \end{pmatrix}$$

since it has rank $=1$ it has a multiplicity ...

$n-1$ has multiplicity 1

$$\mathbb{L} - nI = \frac{1}{n-1} \begin{pmatrix} -1 & -1 & -1 & \cdots & -1 \\ -1 & \ddots & & & \vdots \\ -1 & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \vdots & & & & -1 \\ -1 & & & & -1 \end{pmatrix}$$

$\det(L - nI) = 0$

$\det(\mathbb{1}^{n \times n})(-1) = 0(-1) = 0$

since it has rank $=1$ it has a multiplicity of $(n-1)$

According to proposition 22:
given that there is 1 connected component we know there is 1 eigenvalue 0 with $EV = \mathbb{1}$

$$D = \begin{array}{c} \\ n\{ \\ m\{ \end{array} \overbrace{\begin{bmatrix} m & & & & & \\ & \ddots & & & & \\ & & m & & & \\ & & & n & & \\ & & & & \ddots & \\ & & & & & n \end{bmatrix}}^{\overbrace{\quad\quad}^{n}\ \overbrace{\quad\quad}^{m}}$$

$$A = \begin{array}{c} n\{ \\ m\{ \end{array} \begin{bmatrix} \overbrace{0 \cdots 0}^{n} & \overbrace{1 \cdots 1}^{m} \\ 1 \cdots 1 & 0 \cdots 0 \end{bmatrix}$$

$$L = D^{-1/2}(D-A)D^{-1/2}$$

$$L = D^{-1/2}(D-A)D^{-1/2}$$

$$D-A = \begin{array}{cc} & \begin{array}{cc} \overbrace{\quad\quad\quad}^{n} & \overbrace{\quad\quad\quad}^{m} \end{array} \\ \begin{array}{c} n \\ \\ m \end{array} & \begin{bmatrix} m & & & -1 & \cdots & -1 \\ & \ddots & & \vdots & & \vdots \\ & & m & -1 & \cdots & -1 \\ -1 & \cdots & -1 & n & & \\ \vdots & & \vdots & & \ddots & \\ -1 & \cdots & -1 & & & n \end{bmatrix} \end{array}$$

$$D^{-1/2} = \begin{array}{cc} & \begin{array}{cc} \overbrace{\quad\quad\quad}^{n} & \overbrace{\quad\quad\quad}^{m} \end{array} \\ \begin{array}{c} n \\ \\ m \end{array} & \begin{bmatrix} \frac{1}{\sqrt{m}} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{\sqrt{m}} & & & \\ & & & \frac{1}{\sqrt{n}} & & \\ & & & & \ddots & \\ & & & & & \frac{1}{\sqrt{n}} \end{bmatrix} \end{array}$$

$$\frac{1}{\sqrt{m}} \; \frac{1}{\sqrt{m}} \cdots \frac{1}{\sqrt{m}} \; \frac{1}{\sqrt{n}} \; \frac{1}{\sqrt{n}} \cdots \frac{1}{\sqrt{n}}$$

$$L = \begin{array}{cc} \begin{array}{c} n \left\{ \begin{array}{c} \frac{1}{\sqrt{m}} \\ \frac{1}{\sqrt{m}} \\ \vdots \end{array} \right. \\ m \left\{ \begin{array}{c} \frac{1}{\sqrt{m}} \\ \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{m}} \end{array} \right. \end{array} & \begin{bmatrix} m & & & -1 & \cdots & -1 \\ & m & & -1 & \cdots & -1 \\ & & m & -1 & \cdots & -1 \\ -1 & \cdots & -1 & n & & \\ \vdots & & \vdots & & n & \\ -1 & \cdots & -1 & & & n \end{bmatrix} \end{array}$$

$$tr(L) = n + m$$

$$\sum \lambda_i = 0 + m + n \cancel{-2} + 2$$

$$= m + n$$

$$D = \begin{bmatrix} n-1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & & 0 & & 0 & \\ 1 & 0 & & & 0 & 0 & \\ 1 & 0 & & & & & \\ 1 & & & & & & \end{bmatrix}$$

$$D^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{n-1}} & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

$$L = \begin{array}{cc} & \begin{array}{cccc} \frac{1}{\sqrt{n-1}} & 1 & 1 & 1 \end{array} \\ \begin{array}{c} \frac{1}{\sqrt{n-1}} \\ 1 \\ 1 \\ \vdots \\ 1 \end{array} & \begin{bmatrix} n-1 & -1 & -1 & -1 & \cdots & -1 \\ -1 & 1 & & & & \\ -1 & & 1 & & & \\ \vdots & & & 1 & & \\ -1 & & & & 1 & \\ & & & & & 1 \end{bmatrix} \end{array}$$

$$tr(L) = \frac{n-1}{n-1} + \underbrace{1+1 \cdots +1}_{n-1} = 1+n-1 = n$$

$$\sum \lambda_i = 0 + n-2+2 = n$$