Problem Set Five

Zhixian(Jason) Yu

04/25/2018

1 Theory

Problem 1

Suppose we have a two connected components and the weight matrix W has the following shape:

$$\begin{bmatrix} A_{m\times m} & 0 \\ 0 & B_{n\times n} \end{bmatrix}$$

where A and B are two different components with different shapes. We know that a vector of 1 is an eigenvector of $(D-W)$, where the diagonal elements of D is a sum of row vectors of W, and its associated eigenvalue is 0. It will be easy to see that following two vectors will also be eigenvectors of $(D-W)$ with zero eigenvalues. The two vectors are:

$$v_1 = \begin{bmatrix} \vec{1}_{m\times 1} \\ \vec{0}_{n\times 1} \end{bmatrix} \quad v_2 = \begin{bmatrix} \vec{0}_{m\times 1} \\ \vec{1}_{n\times 1} \end{bmatrix}$$

We can see that:

$$(Dv_1)_i == \begin{cases} D_{ii} = \sum_j w_{ij} & i <= m \\ 0 & i > m \end{cases}$$

and:

$$(Wv_1)_i == \begin{cases} \sum_j^m w_{ij} = \sum_j w_{ij} & i <= m \\ 0 & i > m \end{cases}$$

Therefore $(D-W)v_1 = 0v_1$. This idea can be applied to multiple connected components as well.

Problem 2

Because $M = (W-W)^T(I-W)$, M is a positive semi-definite matrix, and all the eigenvalues of M are non-negative.

Problem 3

a) Because x is a single vector, we have:

$$\begin{aligned}
\|x - \sum_{j\in N} w_j x_j\|^2 &= (x - \sum_{j\in N} w_j x_j)^T (x - \sum_{j\in N} w_j x_j) \\
&= (x^T - \sum_{j\in N} w_j x_j^T)(x - \sum_{j\in N} w_j x_j) \\
&= x^T x - \sum_{j\in N} w_j x^T x_j - \sum_{j\in N} w_j x_j^T x + \sum_{j\in N} w_j x_j^T \sum_{j\in N} w_j x_j \\
&= x^T x - \sum_{j\in N} w_j x^T x_j - \sum_{j\in N} w_j x_j^T x + \sum_{j\in N} w_j x_j^T \sum_{k\in N} w_k x_k \\
&= x^T x - \sum_{j\in N} w_j x^T x_j - \sum_{j\in N} w_j x_j^T x + \sum_{k\in N} \sum_{j\in N} w_j x_j^T (w_k x_k) \\
&= x^T x - \sum_{k\in N} w_k x^T x_k - \sum_{j\in N} w_j x_j^T x + \sum_{jk} w_j w_k x_j^T x_k
\end{aligned}$$

Because $\sum_j w_j = 1$, we have $\sum_k w_k = 1$ and $\sum_{jk} w_j w_k = 1$. Therefore, the above equation becomes:

$$\begin{aligned}
\|x - \sum_{j \in N} w_j x_j\|^2 &= x^T x - \sum_{k \in N} w_k x^T x_k - \sum_{j \in N} w_j x_j^T x + \sum_{jk} w_j w_k x_j^T x_k \\
&= \sum_{jk} w_j w_k x^T x - \sum_j w_j \sum_{k \in N} w_k x^T x_k - \sum_k w_k \sum_{j \in N} w_j x_j^T x + \sum_{jk} w_j w_k x_j^T x_k \\
&= \sum_{jk} w_j w_k x^T x - \sum_{jk} w_j w_k x^T x_k - \sum_{jk} w_j w_k x_j^T x + \sum_{jk} w_j w_k x_j^T x_k \\
&= \sum_{jk} w_j w_k (x^T x - x^T x_k - x_j^T x + x_j^T x_k) \\
&= \sum_{jk} w_j w_k (x - x_j)^T (x - x_k)
\end{aligned}$$

Thus,

$$E(w) = \sum_{jk} w_j w_k C_{jk}$$

where $C_{jk} = (x - x_j)^T (x - x_k)$.

b) Our goal is to minimize $\sum_{jk} w_j w_k C_{jk}$ subject to $\sum_j w_j = 1$. We can apply the method of Lagrange multipliers, and get:

$$L = \sum_{jk} w_j w_k C_{jk} - \lambda(\sum_j w_j - 1)$$

If we take the partial derivative of L over w_j, we get:

$$\begin{aligned}
\frac{\partial L}{\partial w_j} &= 2 w_j C_{jj} + \sum_{k \neq j} w_k C_{jk} + \sum_{k \neq j} w_k C_{kj} - \lambda \\
&= \sum_k w_k C_{jk} + \sum_k w_k C_{kj} - \lambda \\
&= 2 \sum_k w_k C_{jk} - \lambda
\end{aligned}$$

Let $\frac{\partial L}{\partial w_j} = 0$, so:

$$\sum_k w_k C_{jk} = \frac{\lambda}{2}$$

The above equation holds for any reasonable j, therefore:

$$w^T C = \frac{\lambda}{2} \vec{1}^T$$

Because C is a symmetric matrix, we can see that $w = C^{-T} \frac{\lambda}{2} \vec{1} = \frac{\lambda}{2} C^{-1} \vec{1}$. Expanding this leads to the following:

$$w_l = \frac{\lambda}{2} \sum_m C_{lm}^{-1} \tag{1}$$

Because $\sum_l w_l = 1$, we have:

$$\begin{aligned}\sum_l w_l &= \sum_l \frac{\lambda}{2} \sum_m C^{-1}_{lm} \\ &= \frac{\lambda}{2} \sum_{lm} C^{-1}_{lm} \\ &= 1\end{aligned}$$

Therefore $\lambda = \frac{2}{\sum_{lm} C^{-1}_{lm}}$. Substitute the λ from equation 1 with changes of some dummy indexes, we have:

$$w_j = \frac{\sum_k C^{-1}_{jk}}{\sum_{lm} C^{-1}_{lm}}$$

c) Because $M = (I - W)^T (I - W)$, we have:

$$\begin{aligned}M_{ij} &= \sum_k (I - W)^T_{ik} (I - W)_{kj} \\ &= \sum_k (I^T_{ik} - W^T_{ik})(I_{kj} - W_{kj}) \\ &= \sum_k (I_{ik} - W_{ki})(I_{kj} - W_{kj}) \\ &= \sum_k I_{ik} I_{kj} - \sum_k I_{ik} W_{kj} - \sum_k W_{ki} I_{kj} + \sum_k W_{ki} W_{kj} \\ &= \delta_{ij} - W_{ij} - W_{ji} + \sum_k W_{ki} W_{kj}\end{aligned}$$

d)

$$\begin{aligned}\sum_i \left\| y_i - \sum_j W_{ij} y_j \right\|^2 &= \sum_i (y_i - \sum_j W_{ij} y_j)^T (y_i - \sum_j W_{ij} y_j) \\ &= \sum_i (y_i^T - \sum_j W_{ij} y_j^T)(y_i - \sum_j W_{ij} y_j) \\ &= \sum_i y_i^T y_i - \sum_i y_i^T \sum_j W_{ij} y_j - \sum_i \sum_j W_{ij} y_j^T y_i + \sum_i (\sum_j W_{ij} y_j^T \sum_j W_{ij} y_j) \\ &= \sum_i y_i^T y_i - \sum_i \sum_j W_{ij} y_i^T y_j - \sum_i \sum_j W_{ij} y_j^T y_i + \sum_i \sum_k \sum_j W_{ik} y_k^T W_{ij} y_j\end{aligned}$$

We can switch the indexes i and j in the term $\sum_i \sum_j W_{ij} y_j^T y_i$, switch the indexes i and k in the term $\sum_i \sum_k \sum_j W_{ik} y_k^T W_{ij} y_j$, and we get:

$$\begin{aligned}\sum_i \left\| y_i - \sum_j W_{ij} y_j \right\|^2 &= \sum_i \sum_j \delta_{ij} y_i^T y_j - \sum_i \sum_j W_{ij} y_i^T y_j - \sum_j \sum_i W_{ji} y_i^T y_j + \sum_k \sum_i \sum_j W_{ki} y_i^T W_{kj} y_j \\ &= \sum_i \sum_j \delta_{ij} y_i^T y_j - \sum_i \sum_j W_{ij} y_i^T y_j - \sum_i \sum_j W_{ji} y_i^T y_j + \sum_i \sum_j \sum_k W_{ki} W_{kj} y_i^T y_j \\ &= \sum_{i,j} (\delta_{ij} - W_{ij} - W_{ji} + \sum_k W_{ki} y_i^T W_{kj}) y_i^T y_j \\ &= \sum_{i,j} M_{ij} y_i^T y_j\end{aligned}$$

e)

$$\begin{aligned} trace(YMY^T) &= trace(MY^TY) \\ &= \sum_j \sum_i M_{ji}(Y^TY)_{ij} \\ &= \sum_i \sum_j M_{ji} y_i^T y_j \end{aligned}$$

Because $M^T = (I-W)^T(i-W) = M$, M is a symmetric matrix, and $M_{ji} = M_{ij}$. Therefore:

$$\begin{aligned} trace(YMY^T) &= \sum_i \sum_j M_{ij} y_i^T y_j \\ &= \sum_i \| y_i - \sum_j W_{ij} y_j \|^2 \end{aligned}$$

f) We can apply the method of Lagrange multipliers, and get:

$$\begin{aligned} L &= trace(YMY^T) - \sum_{ij} \lambda_{ij}(y_i^T y_j - \delta_{ij}) \\ &= \sum_i \sum_j M_{ij} y_i^T y_j - \sum_{ij} \lambda_{ij}(y_i^T y_j - \delta_{ij}) \end{aligned}$$

If we take the partial derivative of L over y_k, we get:

$$\begin{aligned} \frac{\partial L}{\partial y_k} &= \sum_{j \neq k} M_{kj} y_j + 2M_{kk} y_k + \sum_{j \neq k} M_{jk} y_j - \sum_{j \neq k} \lambda_{jk} y_j - 2\lambda_{kk} y_k - \sum_{j \neq k} \lambda_{kj} y_j \\ &= 2 \sum_j M_{jk} y_j - (\sum_j \lambda_{jk} + \sum_j \lambda_{jk}) y_j \end{aligned}$$

Let $\frac{\partial L}{\partial y_k} = 0$, we have:

$$\sum_j M_{jk} y_j = \frac{1}{2}(\sum_j \lambda_{jk} + \sum_j \lambda_{jk}) y_j$$

Since $\frac{1}{2}(\sum_j \lambda_{jk} + \sum_j \lambda_{jk})$ is a scalar, we can let $\nu = \frac{1}{2}(\sum_j \lambda_{jk} + \sum_j \lambda_{jk})$, and we have:

$$My_i = \nu y_i$$

This is the eigenvector problem.

2 Programming

The following code sets up the environment and import packages.

```
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
import cv2
import random
```

Problem 1