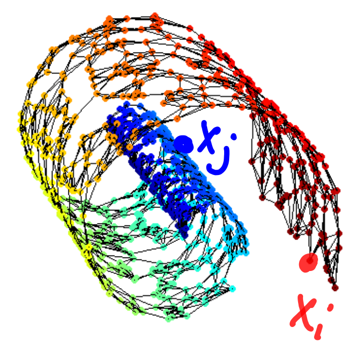
Lecture 14: Intro to Spectral Graph Theory

Last time we discussed the intuition behind the use of graphical models & the underlying manifold assumption.

We now want to explore how such intuition & more broadly, graph models, can be exploited to design algorithms.



Our starting point is a gentle intro to spectral graph theory.

- F. Chung, "Spectral Graph Theory", 1997

14.1 Graph Laplacian Matrices

Suppose we are given a data set $X \subset \mathcal{M} \subseteq \mathbb{R}^d$ $X = \{\underline{x}_1, \dots, \underline{x}_n\}$ last time we saw a simple approach for constructing a graph on this data

$$G = \{X, W\}$$

- Proximity Graphs

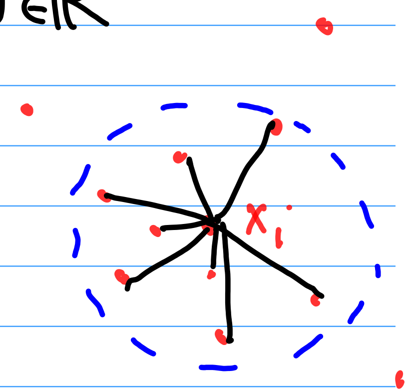
Step1) Pick a weight function/Kernel

$\eta : [0,\infty) \longmapsto [0,\infty)$ non-decreasing

e.g. $\eta(t) = \begin{cases} 1 & \text{if } t \leq r \\ 0 & \text{if } t > r \end{cases}$

Step2) Build weight matrix $W \in \mathbb{R}^{n \times n}$

$$W_{ij} = \eta(\|\underline{x}_i - \underline{x}_j\|)$$



- K-nearest neighbor K-NN

Step1) Pick $K \in \mathbb{N}, K \geq 1$

Step2) For each $\underline{x}_i \in X$ find its K-nearest neighbors $N_K(\underline{x}_i)$

Not symm.

Step3) Build weight matrix $W \in \mathbb{R}^{n \times n}$

$$W_{ij} = \begin{cases} \eta(\|\underline{x}_i - \underline{x}_j\|) & \text{if } \underline{x}_j \in N_K(\underline{x}_i) \\ 0 & \text{otherwise} \end{cases}$$

Step3*) Build weight mat. $W \in \mathbb{R}^{n \times n}$

$$W_{ij} = \begin{cases} \eta(\|\underline{x}_i - \underline{x}_j\|) & \text{if } \underline{x}_j \in N_K(\underline{x}_i) \text{ or } \underline{x}_i \in N_K(\underline{x}_j) \\ 0 & \text{otherwise} \end{cases}$$

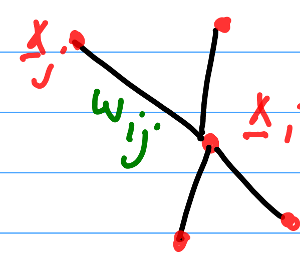
• Proximity graphs are conceptually simpler & easier to analyze.

• K-NN graphs are computationally desirable as their weight matrix W is sparse! ($\approx K$ non-zero entries in each row)

• Note: One can consider many other constructions.

Once we have constructed the graph, ie, the W matrix we can immediately define a so-called graph Laplacian matrix on it.

first, define the degree of the node/vertex $\underline{x}_i$ as



$$d_i = \sum_{j=1}^{n} w_{ij} , \quad i=1,\dots,N$$

This is simply the sum of the rows of W

$$\underline{d} = (d_1, \dots, d_n)^T = W\mathbb{1}, \qquad \mathbb{1} = (1,\dots,1)^T \in \mathbb{R}^n$$

this is also the sum of all weights w_{ij} of edges that are connected to $\underline{x}_i$.

Intuitively, high degree nodes are more influential since they connect to more nodes or have higher weight connections. So the degree encodes the importance of each node.

With the degrees at hand we define the degree matrix

$$\mathbb{R}^{n\times n} \ni D = \mathrm{diag}(\underline{d}), \quad D = \begin{bmatrix} d_1 & & & \emptyset \\ & d_2 & & \\ & & \ddots & \\ \emptyset & & & d_n \end{bmatrix}$$

& in turn define the unnormalized graph Laplacian matrix/operator

$$L := D - W \in \mathbb{R}^{n\times n}$$

We can define other types of graph Laplacians as well. The two most useful ones are

- normalized Graph Laplacian (GL)

(sym.) $$D^{-1/2}(D-W)D^{-1/2} = I - D^{-1/2} W D^{-1/2}$$

• Random walk GL

(non-sym.) $D^{-1}(D-W) = I - D^{-1}W$

14.2 Spectrum of GLs

For simplicity let we will focus on unnormalized Laplacians for now.

Observe that $D-W$ is symm. Also that

$\underline{f} \in \mathbb{R}^n \qquad \underline{f}^T L \underline{f} = \underline{f}^T (D-W)\underline{f} = \underline{f}^T D \underline{f} - \underline{f}^T W \underline{f}$

$$= \sum_{i=1}^{n} d_i f_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} f_i f_j$$

$$= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{ij}\right) f_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} f_i f_j$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(f_i - f_j)^2$$

Thus, L is PDS! But there is more.

$L = D - W \Rightarrow L\mathbb{1} = D\mathbb{1} - W\mathbb{1}$

$$= \underline{d} - \underline{d} = \underline{0}$$

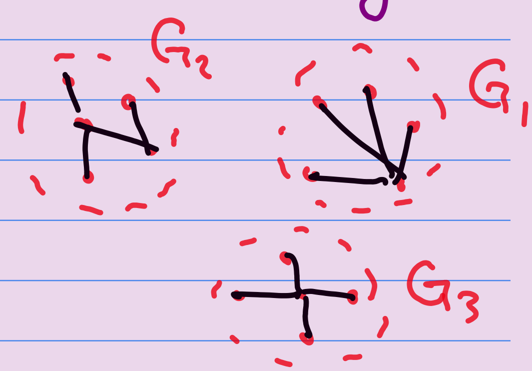
In otherwords, constant vectors belong to the null space of L!

But there is more

<u>Def^n</u> We say the graph G is pathwise connected if for all pairs $\underline{x}_i, \underline{x}_j \in X$ there exists a path (ie seq. of edges) from $\underline{x}_i$ to $\underline{x}_j$.

Observe: if G is NOT Pathwise connected then, after re-labeling of the vertices we can always write W as a block diag. matrix

$$W = \begin{bmatrix} \boxed{W_1} & & \emptyset \\ \emptyset & \boxed{W_2} & \\ & & \ddots & \\ & & & W_M \end{bmatrix}$$



where W_i are the weights of the i-th connected component.

<u>Thm</u>: $\dim(\mathrm{Null}(L)) = 1$ iff G is Pathwise connected.

Idea of proof: suppose it isn't connected & has at least two connected components. Then show that $\dim(\mathrm{Null}(L)) > 1$

This simple fact is the underlying principle idea of the famous spectral clustering alg.

Suppose G has two connected components

$$G = G_1 \cup G_2, \quad G_1 = \{X_1, W_1\}, G_2 = \{X_2, W_2\}$$

Then we can alway re-order pts such that

$$W = \begin{bmatrix} W_1 & \emptyset \\ \emptyset & W_2 \end{bmatrix}$$

We already know that $L\mathbb{1} = \emptyset$ but also since

$$L = D - W = \begin{bmatrix} D_1 & \emptyset \\ \emptyset & D_2 \end{bmatrix} - \begin{bmatrix} W_1 & \emptyset \\ \emptyset & W_2 \end{bmatrix} = \begin{bmatrix} L_1 & \emptyset \\ \emptyset & L_2 \end{bmatrix}$$

where L_1 & L_2 are GL's of G_1 & G_2. So, we can see that

$$\begin{bmatrix} L_1 & \emptyset \\ \emptyset & L_2 \end{bmatrix} \begin{bmatrix} \mathbb{1}_1 \\ \emptyset \end{bmatrix} = \begin{bmatrix} L_1 & \emptyset \\ \emptyset & L_2 \end{bmatrix} \begin{bmatrix} \emptyset \\ \mathbb{1}_2 \end{bmatrix} = \emptyset$$

since $\begin{bmatrix} \mathbb{1}_1 \\ \mathbb{1}_2 \end{bmatrix}, \begin{bmatrix} \mathbb{1}_1 \\ -\mathbb{1}_2 \end{bmatrix}$ are lin. indep. we infer $\dim(Null(L)) \geq 1$ in fact we can show $\dim(Null(L)) = 2$ if G_1 & G_2 are themselves connected.

We just made an important discovery! if G has M-connected components. Then $Nullity(L) = M$! Thus, we simply need to count the # of zero eigenvalues of L to find the # of clusters!

$$\text{Also, } Null(G) = \text{span}\{\mathbb{1}_{G_i}\}_{i=1}^{M}$$

where $\mathbb{1}_{G_i}$ are the indicator vectors of the connected components G_i