Optimization and Computational Linear Algebra for Data Science
Lecture 8: Graphs and Linear Algebra

Léo Miolane · leo.miolane@gmail.com

June 19, 2019

Warning: *This material is not meant to be lecture notes. It only gathers the main concepts and results from the lecture, without any additional explanation, motivation, examples, figures...*

1 Graphs

We start by a formal definition of a (simple non-oriented) graph:

Definition 1.1 (*Graph*)

A graph G is defined as a pair V_G, E_G where $V = V_G$ is the set of vertices of G and $E = E_G$ is the set of edges of G which is a subset of $V \times V$. Two vertices i, j are connected by an edge if $\{i, j\} \in E$. In such case we write $i \sim j$ and say that i and j are neighboors.

Definition 1.2

The degree of a node $i \in V$ is the number of its neighboors.

In this lecture we will only consider finite graphs, where V is finite. We let $n = \#V$. One can assume (up to renaming the vertices) that $V = \{1, \dots, n\}$.

Definition 1.3

We define the adjacency matrix $A \in \mathbb{R}^{n \times n}$ of the graph G by

$$A_{i,j} = \begin{cases} 1 & \text{if } i \sim j \\ 0 & \text{otherwise.} \end{cases}$$

The degree matrix of G is defined by $D = \mathrm{Diag}(\deg(1), \dots, \deg(n))$.

Notice that A is a symmetric matrix.

2 Graph Laplacian

Definition 2.1 (*Graph Laplacian*)

The Laplacian matrix of G is defined as

$$L = D - A.$$

Proposition 2.1

The matrix L satisfies the following properties:

1. *L is symmetric and positive semi-definite.*
2. *The smallest eigenvalue of L is 0, the corresponding eigenvector is the constant one vector $(1, 1, \dots, 1)$.*
3. *L has n non-negative eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.*

Proposition 2.2

The graph G is connected if and only if $\lambda_2 > 0$.

More generally, one can show that the multiplicity of the eigenvalue 0 of L (i.e. the number of i such that $\lambda_i = 0$) is equal to the number of connected components of L.

3 Spectral clustering with the graph Laplacian

Proposition 3.1

Assume that G is connected. Let v_2 be an eigenvector associated to λ_2, the second smallest eigenvalue of L. Let

$$W = \{i \mid v_2(i) \geq 0\}.$$

Then the subgraph induced by W is connected.

This indicates that

$$\operatorname{cut}(G, v) = \sum_{i \sim j} \mathbb{1}(v_i \neq v_i) = \frac{1}{2} \sum_{i,j} A_{i,j}(1 - v_i v_j)$$