1 Graph theoretical preliminaries

By a *network*, we will refer to a structure composed by individual elements, or actors, and by interactions, or connections, between these elements. For example, the world wide web is a network: the elements are the web pages and the connections are the hyperlinks that point from one web page to another. A slightly more complex example is the network made by the behaviour of the reindeer in a herd in a given day: each reindeer is an actor, and we place a weighted interaction between any two reindeer: the weight is n if they come closer than a given threshold (say, 1 meter) n times in a day. For more examples and discussions see the recommended books.

1.1 Basic definitions

The natural mathematical object to model a network is a *graph*.

Definition 1.1. *A* graph $G = (V, E)$ *is an ordered pair of sets where:*

- V *is the set of* vertices*;*
- $E \subseteq V \times V$ *is the set of* edges*.*

The set of the vertices and the set of the edges of a given graph G will be denoted by $V(G)$ and $E(G)$, respectively. Frequently used synonyms of "vertex" include "node" or "point", and common synonyms of "edge" include "link" or "line". And edge has the form (x, y) for some $x, y \in V$; given the edge (x, y), we say that the vertices x and y are its *endvertices* (or endpoints, or endnodes).

Graphs can either be *undirected* or *directed*. An undirected graph G is such that $(x, y) \in E(G) \Leftrightarrow (y, x)$. Hence, in an undirected graph, an edge can (and often is) be represented by the *unordered* pair of its endpoints x, y. Conversely, in a directed graph it may happen that $(x, y) \in E(G)$ but $(y, x) \notin E(G)$. When both (x, y) and (y, x) are in $E(G)$ for a directed graph G, we say that the edge (x, y) is *reciprocal*. Below, we will often somewhat simplify our notation by using the shorthand notation xy to denote the edge (x, y). We say that the vertex x is the *tail* (or origin) and that the vertex y is the *head* (or destination) of the edge xy.

A *loop* is an edge of the form (x, x), i.e., an edge whose head and tail coincide. If a graph G is such that $e \in E(G) \Rightarrow \nexists x \in V(G) : e = (x, x)$, we say that G is *without loops*; otherwise, G is with loops.

A *weighted graph* is a graph $G = (V, E)$ together with a function $\omega : E \to \mathbb{R}_+$; the positive real number $\omega(e)$ is called the *weight* of the edge $e \in E$. An *unweighted graph* is a graph without such a function, or it can equivalently be seen as a weighted graph together with the weight function constantly equal to 1, i.e., $\omega(e) = 1 \ \forall \ e \in E$.

A *simple graph* is a graph which is unweighted, undirected, and without loops.

A *walk of length* ℓ is a sequence of nodes $i_1, \ldots, i_{\ell+1}$ such that $i_k i_{k+1}$ exists for $k = 1, \ldots, \ell$; within a walk, both nodes and edges may generally appear more than once. A *path of length* ℓ is a walk of length ℓ such that either all the nodes are distinct (open path) or all the nodes are distinct except for $i_1 = i_{\ell+1}$ (closed path). A closed path of length $\ell \geq 3$ is called a *cycle of length* ℓ.

Example 1.1. *Let us consider the graph* $G = (V, E)$ *given by* $V(G) = \{1, 2, 3, 4\}$ *and* $E(G) = \{13, 31, 14, 41, 34, 43, 24, 42\}$*. Then* G *is a simple graph: it does not have loops, no weights were defined, and it is undirected since, for every edge in* $E(G)$ *the reciprocal is also in* $E(G)$*. An example of a walk in* G *is* $1, 3, 1, 4, 2$*: however, this walk is not a path as the node* 1 *appears twice, and in one occasion not in an extremal position. An example of an open path in* G *is* $2, 4, 1$*. An example of a cycle of length* 3 *in* G *is* $1, 4, 3, 1$*.*

Finally, we recall that an undirected graph $G(V, E)$ is *bipartite* if there exists a partition $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$ such that, for all $ij, ji \in E$ then either $i \in V_1, j \in V_2$ or $i \in V_2, j \in V_1$.

1.2 Exercises

1. Let $G = (V, E)$ be the Petersen graph, which is defined by $V(G) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 0\}$ and $E(G) = \{12, 15, 16, 21, 23, 27, 32, 34, 38, 43, 45, 49, 51, 54, 50, 61, 68, 69, 72, \ 79, 70, 83, 86, 80, 94, 96, 97, 05, 07, 08\}$. Is G simple? Give two distinct examples of cycles of length 5 in G.

2 Algebraic graph theory

In this section, we define some important matrices that are associated with a graph. First, though, we need to introduce an important equivalence class on graphs.

Definition 2.1. *We say that two graphs* $G_1 = (V_1, E_1)$ *and* $G_2 = (V_2, E_2)$ *are* isomorphic*, and we write* $G_1 \simeq G_2$*, if there is a bijection* $f : V_1 \to V_2$ *such that*

$$\forall x, y \in V_1 \qquad xy \in E_1 \Leftrightarrow f(x)f(y) \in E_2.$$

The bijection f is called a graph isomorphism; if, moreover, f is a graph isomorphism of G into itself (i.e., a reordering of the vertices) then it called an automorphism of G. For all practical purposes of this course, we will identify graphs that are isomorphic to each other (or, in other words, we can safely work with the equivalence class of all graphs isomorphic to a given one). For this reason, we will often informally apply automorphisms to graphs to simplify proofs and arguments.

The first important matrix that we associate with a graph is called the *adjacency matrix.*

Definition 2.2. *The* adjacency matrix *of an unweighted graph* $G = (V, E)$*, with* $V(G) = \{1, \ldots, n\} =: [n]$*, is an* $n \times n$ *matrix* $A(G)$ *with entries*

$$A_{ij} = \begin{cases} 1 \text{ if } ij \in E(G); \\ 0 \text{ otherwise.} \end{cases}$$

The adjacency matrix of a weighted graph $G = (V, E)$ *with* $V(G) = [n]$ *the adjacency matrix is the* $n \times n$ *matrix* $A(G)$ *with entries*

$$A_{ij} = \begin{cases} \omega(ij) \text{ if } ij \in E(G); \\ 0 \text{ otherwise.} \end{cases}$$

The adjacency matrix has a natural and useful combinatorial property related to walks.

Theorem 2.1. *Let* G *be an unweighted graph. Then,* A^k_{ij} *is the number of walks of length* k *from node* i *to node* j*.*

Proof. We give a proof by induction on k. A walk of length 1 is simply a sequence of 2 nodes i, j such that $ij \in E(G) \Leftarrow A_{ij} = 1$. Clearly, given any two nodes i, j, the number of walks of length 1 from i to j is either 1 (iff $A_{ij} = 1$) or 0 (iff $A_{ij} = 0$).

Now denote $\#_w(i, j, k)$ the number of walks of lenght k from i to j, and let n be the number of vertices in G. Assume that $\#_w(i, \ell, k-1) = A_{i\ell}^{k-1}$. Then,

$$\#_w(i, j, k) = \sum_{\ell : \ell j \in V(G)} \#_w(i, \ell, k-1) = \sum_{\ell=1}^{n} A_{i\ell}^{k-1} A_{\ell j} = A_{ij}^{k}.$$

□

Recall that a *permutation matrix* P is a square matrix having precisely one element equal to 1 in each row and column, and every other element equal to 0. Permutation matrices are orthogonal, i.e., they satisfy the equation $P^T P = I$. (The proof is left as an exercise).

Theorem 2.2. *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs having adjacency matrices, respectively, A_1 and A_2. $G_1 \cong G_2$ if and only if there exists a permutation matrix P such that $A_1 = P^T A_2 P$.*

Proof. Suppose first that the two adjacency matrices are permutation similar, i.e., $A_1 = P^T A_2 P$ for some permutation matrix P. Then, A_1 and A_2 have the same size, say, n. Without loss of generality (up to relabelling the nodes if necessary) we may therefore assume that $V_1 = V_2 = [n]$. For all $i = 1, \ldots, n$, define the function $f(i) = j$ if and only if the ith column of P is e_j (the jth vector in the canonical basis of $\mathbb{R}^n$). Then, f is a bijection (proof as an exercise!). Moreover, by definition of f, it is clear that $Pe_i = e_{f(i)}$.

Suppose now that $ij \in E_1$, which by definition of adjacency matrix is equivalent to $(A_1)_{ij} = 1$. Then,

$$1 = e_i^T A_1 e_j = (Pe_i)^T A_2 (Pe_j) = e_{f(i)}^T A_2 e_{f(j)} = (A_2)_{f(i), f(j)},$$

and hence, $f(i)f(j) \in E_2$. Therefore the two graphs are isomorphic.

Conversely, suppose that the two graphs are isomorphic. Then, in particular A_1 and A_2 have the same size, say, n, and we can again assume $V_1 = V_2 = [n]$. Let f be the bijection associated to the isomorphism and *define* P to satisfy $Pe_i = e_{f(i)}$ for all i. Then,

$$(P^T A_2 P)_{ij} = (Pe_i)^T A_2 (Pe_j) = (A_2)_{f(i),f(j)} = A_{ij} = \begin{cases} 1 \text{ if } f(i)f(j) \in E_2 \Leftrightarrow ij \in E_1; \\ 0 \text{ otherwise.} \end{cases}$$

It follows that $P^T A_2 P = A_1$. □

A consequence of Theorem 2.2 is that the adjacency matrices of isomorphic graphs are (permutation) similar. Therefore, they all share the same eigenvalues. The converse, however, is not true – two graphs may have isospectral adjacency matrices without being isomorphic to each other.

Note, moreover, that a graph automorphism (or, equivalently by Theorem 2.2, a permutation similarity of adjacency matrices) amounts to a *relabelling* of the vertices, i.e., assigning to them arbitrary values in $[n]$.

The following definition has to be interpreted considering all graphs as directed, i.e., we consider all reciprocal edges (or all undirected edges for an undirected graph) as a pair of directed edges.

Definition 2.3. *The* incidence matrix *of a graph without loops $G = (V, E)$, with $V(G) = [n]$ and $\#E(G) = m$, is an $n \times m$ matrix B with entries*

$$B_{ij} = \begin{cases} -1 \text{ if the tail of edge } j \text{ is node } i; \\ 1 \text{ if the head of edge } j \text{ is node } i; \\ 0 \text{ otherwise.} \end{cases}$$

Definition 2.4. *Let $G = (V, E)$ be a simple graph. The* degree *of a vertex $v \in V(G)$ is the number of undirected edges having v as an endvertex.*

Note that in the definition of degree we are counting undirected edges, unlike for example in the definition of incidence matrix when we were considering directed edges. Therefore, for example, if the vertices having endvertex i are precisely ij, ji, ik and ki, then the degree of i is 2 (and not 4).

Definition 2.5. *Let $G = (V, E)$ be a simple graph. The* graph Laplacian *is the $n \times n$ matrix defined as $L = \Delta - A$, where $A = A(G)$ is the adjacency matrix of the graph G and Δ is a diagonal matrix such that Δ_{ii} is the degree of node i. Moreover, the* normalized Laplacian *is the matrix $\Delta^{-1/2} L \Delta^{-1/2} = I - \Delta^{-1/2} A \Delta^{-1/2}$, where $\Delta^{-1/2}$ is the diagonal matrix whose ith diagonal entry is $(\deg i)^{-1/2}$.*

Theorem 2.3. *For a simple graph, the graph Laplacian and the incidence matrix are related by the formula* $L = \frac{1}{2}BB^T$.

Proof. Suppose that the graph has m edges and n nodes. Observe that $(BB^T)_{ij} = \sum_{k=1}^{m} B_{ik}B_{jk}$. If $i \neq j$, then this sum is equal to -2 if $ij, ji \in E(G)$ (recall that a simple graph is undirected) or it is equal to 0 otherwise. If $i = j$, then this sum counts the number of directed edges having node i as an endpoint, i.e., twice the number of undirected edges having node i as an endpoint, i.e., twice the value of Δ_{ii}. Hence, $BB^T = 2\Delta - 2A = 2L$. $\square$