Lesson 3

Algebraic graph theory

Sergio Barbarossa
Basic notions

Definition: A directed graph (or digraph) \( G = \{V, E\} \) is composed by a set of vertices
and a set of edges \( E \subseteq V \times V \):

We adopt the convention that \( e_{ij} \triangleq (v_i, v_j) \in E \) means that
the information flows from node j to node i.

A digraph is weighted if a positive weight (not necessarily 1)
is associated to each edge.

We assume that the graph is acyclic, i.e. \( a_{ii} = 0 \).

For any node \( v_i \), we define the neighbors of \( v_i \) as
\[
N_i \triangleq \{j = 1, \ldots, N : e_{ij} = (v_i, v_j) \in E\}.
\]
Basic notions

The in-degree and out-degree of node $v_i$ are defined as

$$\text{deg}_{in}(v_i) := \sum_{j=1}^{N} a_{ij}, \quad \text{and} \quad \text{deg}_{out}(v_i) := \sum_{j=1}^{N} a_{ji}$$

For undirected graphs

$$\text{deg}_{in}(v_i) = \text{deg}_{out}(v_i)$$

Classes of digraphs:

1. **Balanced digraph**: The node $v_i$ of a digraph is balanced if and only if its in-degree and out-degree are equal.

A graph is balanced if and only if all its nodes are balanced

$$\sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji}, \quad \forall i = 1, \ldots, N$$
2. **Path**: A **strong path** (or directed chain) is a sequence of distinct nodes $(v_0, \ldots, v_q) \in \mathcal{V}$ such that $(v_i, v_{i-1}) \in \mathcal{E}$.

If $v_0 \equiv v_q$, the path forms a cycle.

A **weak path** is a sequence of $q$ distinct nodes such that either $(v_{i-1}, v_i) \in \mathcal{E}$ or $(v_i, v_{i-1}) \in \mathcal{E}$, $\forall i = 1, \ldots, q - 1$.

3. **Tree**: A digraph with $N$ nodes is a directed tree if it has $N-1$ edges and there exists a distinguished node, called the root node, which can reach all the other nodes by a (unique) strong path.
A directed tree cannot have cycles and every node, except the root, has one and only one incoming edge.

A digraph is a (directed) forest if it consists of one or more directed trees.

A subgraph $G_s = \{V_s, E_s\}$ of a digraph $G = \{V, E\}$ is a directed spanning tree (or a spanning forest), if it is a directed tree (or a directed forest) and it has the same node set as $G$, i.e. $V_s \equiv V$.

A subgraph contains a spanning tree (or a spanning forest) if there exists a subgraph of $G$ that is a directed spanning tree (or a spanning forest).
Connectivity

A digraph is **strongly** connected (SC) if any ordered pair of distinct nodes can be joined by a strong path.

A digraph is **quasi strongly** connected (QSC) if, for every pair of nodes $v_i$ and $v_j$, there exists a node $v_r$ that can reach both nodes by a strong path.

A digraph is **weakly** connected (WC) if any pair of distinct nodes can be joined by a weak path.

In general SC $\Rightarrow$ QSC $\Rightarrow$ WC, but the converse it is not true.
A maximal subgraph which is also SC forms a strongly connected component (SCC)

Any digraph can be partitioned into SCCs, let us say

\[ \mathcal{G}_1 = \{\mathcal{V}_1, \mathcal{E}_1\}, \ldots, \mathcal{G}_K = \{\mathcal{V}_K, \mathcal{E}_K\} \]

A digraph \( \mathcal{G} \) can always be reduced to the corresponding **condensation** digraph \( \mathcal{G}^* = \{\mathcal{V}^*, \mathcal{E}^*\} \) by associating the node set \( \mathcal{V}_i \) of each SCC of \( \mathcal{G} \) to a distinct node \( \nu^*_i \in \mathcal{V}^* \) of the condensation digraph and introducing an edge from \( \nu^*_i \) to \( \nu^*_j \) if and only if there exists some edges from the i-th and the j-th SCC’s
Connectivity

An SCC that is reduced to the root of a directed spanning tree of the condensation digraph is called root SCC (RSCC)

By definition, the condensation digraph has no cycles

There always exists an ordering $v_1^*, \ldots, v_K^*$ of the nodes of the condensation digraph such that all the edges of $\mathcal{G}^*$ are in the form

$$(v_i^*, v_j^*) \in \mathcal{E}^*, \quad \text{with} \quad 1 \leq j < i \leq K,$$

where $v_1^*$ has zero in-degree
Connectivity

The connectivity properties of a digraph can be easily checked looking at the structure of its condensation digraph:

P.1: A graph $\mathcal{G}$ is SC if and only if its condensation graph $\mathcal{G}^*$ is composed by a single node

P.2: A graph $\mathcal{G}$ is QSC if and only if its condensation graph $\mathcal{G}^*$ contains a directed spanning tree

P.3: If $\mathcal{G}$ is WC, then $\mathcal{G}^*$ contains either a spanning tree or a (weakly) connected forest
Connectivity

Examples

Strongly-connected digraph

Quasi-strongly connected digraph

Weakly-connected digraph
Forest with two trees
Algebraic graph theory

The main properties of a graph composed of $N$ vertices and $E$ edges can be studied introducing the following matrices:

**Adjacency matrix**: $A$ is an $N \times N$ matrix whose entry $a_{ij} > 0$ if there is an edge $e_{ij} \in \mathcal{E}$ between nodes $i$ and $j$; otherwise $a_{ij} = 0$.

**Degree matrix**: $D$ is an $N \times N$ diagonal matrix with entries

$$d_{ii} = \sum_{j=1}^{N} a_{ij}$$

**Laplacian**: $L = D - A$
Algebraic graph theory

Some structural properties of a graph resulting from the properties of $A$

P.1 - The number of walks from $u$ to $v$ with length $r$ is $A^r(u, v)$

P.2 - $\text{tr}(A) = 0$

P.3 - $\text{tr}(A^2) = 2E$

If $E$ denotes the number of edges and $T$ the number of triangles

P.3 - $\text{tr}(A^3) = 6T$
Algebraic graph theory

**Incidence matrix:** $B$ is an $N \times E$ matrix with

$B_{i,j} = 1$ if vertex $i$ is the head of edge $j$

$B_{i,j} = -1$ if vertex $i$ is the tail of edge $j$

$B_{i,j} = 0$ otherwise

Property: Let $G$ be a graph with $N$ vertices and $c$ connected components

$\Rightarrow \quad \text{rank}(B) = N - c$

Relationship between graph matrices:

$BB^T = D - A = L$

Rome - February 2010
Algebraic graph theory

The number of connected components is equal to the multiplicity of the null eigenvalue of the Laplacian

Laplacian and quadratic forms

\[ x^T L x = \sum_{(u,v) \in E} (x_u - x_v)^2 \]

The proof is trivial observing that \( x^T L x = x^T B B^T x \) and the entry of \( B^T x \) corresponding to \((u,v)\) is \( \pm (x_u - x_v) \)
Properties of the Laplacian matrix:

i) it is a diagonally dominant matrix
ii) it has zero row sums
iii) it has nonnegative diagonal elements

From i)-iii), invoking Gershgorin's disk theorem, $0$ is an eigenvalue of $L$ with corresponding right eigenvector belonging to $\text{Null}\{L\} \supseteq \text{span}\{1_N\}$ or

$$L \ 1_N = 0_N$$

All the other eigenvalues have positive real part

$$\Rightarrow \text{rank}(L) \leq N - 1$$
A digraph is balanced if and only if $1_N$ is also a left eigenvector of $L$ associated with the zero eigenvalue, i.e.,

$$1_T^T N L = 0_T^T N$$

Connectivity properties derived by the zero eigenvalue

P.1 The multiplicity of the zero eigenvalue of $L$ is equal to the minimum number of directed trees contained in a directed spanning forest of $G$

P.2 The zero eigenvalue of $L$ is simple if and only if $G$ contains a spanning directed tree or, equivalently, it is QSC
As a consequence, if \( G \) is SC, then \( L \) has a simple zero eigenvalue and a positive left-eigenvector associated to the zero eigenvalue.

The Laplacian of a QSC digraph has a simple eigenvalue equal to zero, corresponding to a right eigenvector in \( \text{span}\{1_N\} \).

For undirected graphs, \( \text{rank}(L) = N - 1 \) if and only if \( G \) is connected.

For directed graphs, the “only if” part does not hold.
Algebraic graph theory

Structure of the left-eigenvector $\gamma$ associated to the null eigenvalue of the Laplacian matrix

Thm: Assume that $\mathcal{G}$ is QSC, with $K \geq 1$ strongly connected components (SCC) $\mathcal{G}_1 = \{V_1, E_1\}, \ldots, \mathcal{G}_K = \{V_K, E_K\}$, with $r_i = |V_i|$ ordered so that $\mathcal{G}_1$ coincides with the root SCC of $\mathcal{G}$,

$$\Rightarrow \gamma_i = \begin{cases} > 0, & \text{iff } v_i \in V_1, \\ = 0, & \text{otherwise.} \end{cases}$$
Algebraic graph theory

Sketch of the proof:

There exists an ordering of the vertices, passing through the condensation digraph, such that the Laplacian can be written in block triangular form

\[
L(G) = \begin{pmatrix}
    L_1 & 0 & \ldots & 0 \\
    * & B_2 & \ddots & \vdots \\
    * & * & \ddots & 0 \\
    * & \ldots & * & B_K
\end{pmatrix}
\]

where

\[L_1 = L(G_1)\] is the Laplacian matrix associated to the root SCC

\[B_k = L_k + D_k\] where \[L_k = L(G_k)\] is associated to the k-th SCC

\[D_k\] is a nonnegative diagonal matrix
Algebraic graph theory

The i-th entry of $D_k$ is equal to the sum of the weights associated to the edges outgoing from the nodes in $\{G_1, \ldots, G_{k-1}\}$ and incoming to the i-th node in $G_k$.

Note: Since the graph is QSC, each $D_k$ has at least one positive entry, otherwise the SCC $G_k$ would be decoupled from the other SCCs.

Since each $G_k$ is SC by definition,

$$\Rightarrow \quad rank(L_k) = r_k - 1 \quad \text{and} \quad Null(L_k) = span(1_{r_k})$$

Since $D_k 1_{r_k} \neq 0_{r_k}$,

$$\Rightarrow \quad rank(B_k) = r_k, \quad \forall k = 2, \ldots, K$$
Algebraic graph theory

Partitioning the left eigenvector as

we have

\[ \gamma = [\gamma_{r_1}, \gamma_{r_2}, \ldots, \gamma_{r_K}]^T \]

If, in addition, \( G_1 \) is balanced, then

\[ \gamma_{r_1} > 0_{r_1} \quad \text{and} \quad \gamma_{r_k} = 0_{r_k}, \quad k = 2, \ldots, K \]

\[ \gamma_{r_1} \in span \{1_{r_1}\} \]
References