Graphs – Introduction (theory) – 3

MATH 2740 – Mathematics of Data Science – Lecture 17

Julien Arino

julien.arino@umanitoba.ca

Department of Mathematics @ University of Manitoba

Fall 202X

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis. We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

Outline

Directed graphs

Matrices associated to a graph/digraph

Directed graphs

Matrices associated to a graph/digraph

Directed graphs

- **Degrees in digraphs**
- Walks, paths, etc.
- Connectivity in digraphs
- Orientable graphs

Degree

Let v be a vertex of a digraph $G = (V, A)$

Definition 96 (Outdegree of a vertex)

The number of arcs directed away from a vertex v, in a digraph is called the **outdegree** of v and is written $d_G^+(v)$

Definition 97 (Indegree of a vertex)

The number of arcs directed towards a vertex v, in a digraph is called the **indegree** of v and is written $d_G^-(v)$

Definition 98 (Degree)

For any vertex v in a digraph, the **degree** of v is defined as

$$d_G(v) = d_G^+(v) + d_G^-(v)$$

```r
# Degrees in a digraph: out-, in-, and total degree
deg_out <- degree(G, mode = "out")
deg_in  <- degree(G, mode = "in")
deg_tot <- degree(G, mode = "all")
head(data.frame(v = as.integer(V(G)), out = deg_out, in_deg = deg_in, total
```

```
##   v out in_deg total
## 1 1   2      2     4
## 2 2   2      2     4
## 3 3   2      2     4
## 4 4   2      2     4
## 5 5   3      2     5
## 6 6   2      2     4
```

Theorem 99

For any (di)graph, the sum of the degrees of the vertices equals twice the number of edges (arcs)

Corollary 100

In any (di)graph, the sum of the degrees of the vertices is a nonnegative even integer

Theorem 101

If G is a digraph with vertex set $V(G) = \{v_1, \ldots, v_p\}$ and q arcs, then

$$\sum_{i=1}^{p} d_G^+(v_i) = \sum_{i=1}^{p} d_G^-(v_i) = q$$

```r
# A perfectly symmetric digraph would have every arc reciprocated
recip <- reciprocity(G)  # ratio of mutual connections
recip
```

```
## [1] 0.05128205
```

```r
# Are all edges mutual? (strict symmetry)
all_mutual <- all(is.mutual(G, E(G)))
all_mutual
```

```
## [1] FALSE
```

Definition 102 (Regular digraph)

A digraph G is r-regular if $d_G^+(v) = d_G^-(v) = r$ for all $v \in V(G)$

Symmetric/antisymmetric digraphs

Definition 103 (Symmetric digraph)

Let $G = (V, A)$ be a digraph with associated binary relation R. If R is *symmetric*, the digraph is symmetric

Definition 104 (Anti-symmetric digraph)

Let $G = (V, A)$ be a digraph with associated binary relation R. The digraph G is **anti-symmetric** if

$$xRy \implies y\not{R}x$$

Definition 105 (Symmetric multidigraph)

Let $G = (V, A)$ be a multidigraph. G is symmetric if $\forall x, y \in V(G)$, the number of arcs from x to y equals the number of arcs from y to x

```
# Shortest paths between two vertices
sp_1_to_15 <- shortest_paths(G, from = 1, to = 15, mode = "out")
sp_1_to_15$vpath

## [[1]]
## + 4/20 vertices, from 2a87b22:
## [1]  1  2  5 15

# All simple paths up to a given cutoff (to keep it small)
paths_1_to_6 <- all_simple_paths(G, from = 1, to = 6, mode = "out", cutoff =
length(paths_1_to_6)

## [1] 4

# Number of walks of length 2 via adjacency powers
Mdir <- as.matrix(as_adjacency_matrix(G, sparse = FALSE))
M2 <- Mdir %*% Mdir
M2[1, 6]

## [1] 0
```

Directed graphs

- Degrees in digraphs
- **Walks, paths, etc.**
- Connectivity in digraphs
- Orientable graphs

Walks

Let $G = (V, A)$ be a digraph.

Definition 106 (Directed walk)

A **directed walk** in a digraph G is a non-empty alternating sequence $v_0 a_0 v_1 a_1 v_2 \ldots a_{k-1} v_k$ of vertices and arcs in G such that $a_i = (v_i, v_{i+1})$ for all $i < k$. This walk begins with v_0 and ends with v_k

Definition 107 (Length of a directed walk)

The length of a directed walk is equal to the number of arcs in the directed walk

Definition 108 (Closed walk)

If $v_0 = v_k$, the walk is closed

```r
is_connected(G, mode = "weak")

## [1] TRUE

is_connected(G, mode = "strong")

## [1] FALSE

scc <- components(G, mode = "strong")
scc$no

## [1] 2

table(scc$membership)

##
##  1  2
## 12  8
```

Trails

Let $G = (V, A)$ be a digraph.

Definition 109 (Directed trail)

A directed walk in G in which all arcs are distinct is a **directed trail** in G

Definition 110 (Directed path)

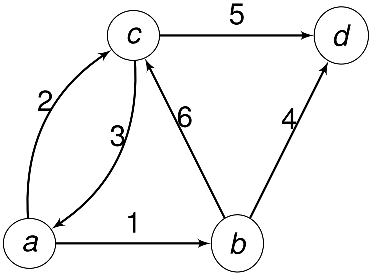
A directed walk in G in which all vertices are distinct is a **directed path** in G

Definition 111 (Directed cycle)

A closed walk is a **directed cycle** if it contains at least three vertices and all its vertices are distinct except for $v_0 = v_k$

```r
# Condensation graph (SCC DAG) via contracting SCCs
scc <- components(G, mode = "strong")
G_cond <- contract(G, scc$membership, vertex.attr.comb = "ignore")
G_cond <- simplify(G_cond, remove.multiple = TRUE, remove.loops = TRUE)

vcount(G_cond); ecount(G_cond)

## [1] 2
## [1] 1

is_dag <- is.dag(G_cond)
is_dag

## [1] TRUE
```

Examples of directed cycles



Cycles:

- $\mu^1 = (1, 6, 2) = [abca]$
- $\mu^2 = (1, 6, 3) = [abca]$
- $\mu^3 = (2, 3) = [aca]$
- $\mu^4 = (1, 4, 5, 2) = [abdca]$
- $\mu^5 = (6, 5, 4) = [acdb]$
- $\mu^6 = (1, 4, 5, 3) = [abdca]$

```r
# Orientation and bridges (undirected case -> orientation)
Gu <- make_ring(8)
length(bridges(Gu))
```

```
## [1] 0
```

```r
G_or <- as.directed(Gu, mode = "arbitrary")  # orient edges
is_connected(G_or, mode = "strong")
```

```
## [1] FALSE
```

Directed graphs

Degrees in digraphs

Walks, paths, etc.

Connectivity in digraphs

Orientable graphs

Definitions

Definition 112 (Underlying graph)

Given a digraph, the undirected graph with each arc replaced by an edge is called the **underlying graph**

Definition 113 (Weakly connected digraph)

If the underlying graph is a connected graph, then the digraph is **weakly connected**

Definition 114 (Strongly connected digraph)

A digraph G is **strongly connected** if for every two distinct vertices u and v of G, there exists a directed path from u to v

Definition 115 (Disconnected digraph)

A digraph is said to be **disconnected** if it is not weakly connected

```r
# Adjacency matrix (directed)
M <- as.matrix(as_adjacency_matrix(G, sparse = FALSE))

# Column sums = indegrees, row sums = outdegrees (directed case)
all.equal(rowSums(M), degree(G, mode = "out"))

## [1] TRUE

all.equal(colSums(M), degree(G, mode = "in"))

## [1] TRUE

# Walks of length 3 from 1 to 15 via M^3
M3 <- M %*% M %*% M
M3[1, 15]

## [1] 2
```

Strong connectedness is an equivalence relation

Denote $x \equiv y$ the relation "$x = y$, or $x \neq y$ and there exists a directed path in G from x to y". $\equiv$ is an equivalence relation since

1. $x \equiv y$ [reflexivity]
2. $x \equiv y \implies y \equiv x$ [symmetry]
3. $x \equiv y, y \equiv z \implies x \equiv z$ [transitivity]

Definition 116 (Connected component of a graph)

Sets of the form

$$A(x_0) = \{x : x \in V, x \equiv x_0\}$$

are equivalence classes. They partition V into strongly connected sub-digraphs of G called **strongly connected components** (or **strong components**) of G

A strong component in G is a maximal strongly connected subdigraph of G

```r
# Distance matrix (shortest path lengths)
D <- distances(G, mode = "out")
D[1, 15]

## [1] 3
```

Theorem 117 (Properties)

Let $G = (V, A)$ be a digraph

- *If G is strongly connected, it has only one strongly connected component*
- *The strongly connected components partition the vertices $V(G)$, with every vertex in exactly one strongly connected component*

```
# Degree matrix and Laplacian
deg_vec <- degree(G, mode = "all")
Dmat <- diag(deg_vec)
L <- laplacian_matrix(G, normalized = FALSE, sparse = FALSE)

dim(Dmat); dim(L)

## [1] 20 20
## [1] 20 20

eigenvals <- eigen(as.matrix(L))$values
head(eigenvals)

## [1] 4.125310+0.0000000i 2.957548+1.4580892i 2.957548-1.4580892i
## [4] 2.598554+0.8441142i 2.598554-0.8441142i 2.172163+0.0000000i
```

Algorithm for determining strongly connected components in $G = (V, A)$

- Determine the strongly connected component $C(v)$ containing the vertex v; if $V - C(v)$ is non-empty, re-do the same operation on the sub-digraph $G' = (V - C(v), A')$
- To determine $C(v)$, the strongly connected component containing v: let v be a vertex of a digraph , which is not already in any strongly connected component
  1. Mark the vertex v with $\pm$
  2. Mark with $+$ all successors (not already marked with $+$) of a vertex marked with $+$
  3. Mark with $-$ all predecessors (not already marked with $-$) of a vertex marked with $-$
  4. Repeat until no more possible marking with $+$ or $-$

  All vertices marked with $\pm$ belong to the same strongly connected component $C(v)$ containing the vertex v

Condensation of a digraph

Definition 118 (Condensation of a digraph)

The condensation G^* of a digraph G is a digraph having as vertices the strongly connected components (SCC) of G and such that there exists an arc in G^* from a SCC C_i to another SCC C_j if there is an arc in G from some vertex of S_i to a vertex of S_j

Definition 119 (Articulation set)

For a connected graph, a set X of vertices is called an **articulation set** (or a **cutset**) if the subgraph of G generated by $V - X$ is not connected

Definition 120 (Stable set)

A set S of vertices is called a **stable set** if no arc joins two distinct vertices in S

Directed graphs

- Degrees in digraphs
- Walks, paths, etc.
- Connectivity in digraphs
- **Orientable graphs**

Orientation

Definition 121 (Orienting a graph)

Given a connected *graph*, we describe the act of assigning a direction to each edge (edge $\rightarrow$ arc) as **orienting the graph**

Definition 122 (Strong orientation)

If the digraph resulting from orienting a graph is strongly connected, the orientation is a **strong orientation**

Orientable graph

Definition 123 (Orientable graph)

A connected graph G is **orientable** if it admits a strong orientation

Theorem 124

A connected graph $G = (V, E)$ is orientable $\iff$ G contains no bridges

(in other words, iff every edge is contained in a cycle)

Matrices associated to a graph/digraph

There are multiple matrices associated to a graph/digraph

The branch of graph theory that studies the properties of matrices derived from graphs and uses of these matrices in determining graph properties is *spectral graph theory*

Graphs greatly simplify some problems in linear algebra and vice versa

Matrices associated to a graph/digraph

- **Adjacency matrices**
- Other matrices associated to a graph/digraph
- Linking graphs and linear algebra

Adjacency matrix (undirected case)

Let $G = (V, E)$ be a graph of order p and size q, with vertices $v_1, \ldots, v_p$ and edges $e_1, \ldots, e_q$

Definition 125 (Adjacency matrix)

The **adjacency matrix** is

$$M_A = M_A(G) = [m_{ij}]$$

is a $p \times p$ matrix in which

$$m_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

Theorem 126 (Adjacency matrix and degree)

The sum of the entries in row i of the adjacency matrix is the degree of v_i in the graph

We often write $A(G)$ and, reciprocally, if A is an adjacency matrix, $G(A)$ the corresponding graph

G undirected $\implies$ $A(G)$ symmetric

$A(G)$ has nonzero diagonal entries if G is not simple

Adjacency matrix (directed case)

Let $G = (V, A)$ be a digraph of order p with vertices $v_1, \ldots, v_p$

Definition 127 (Adjacency matrix)

The **adjacency matrix** $M = M(G) = [m_{ij}]$ is a $p \times p$ matrix in which

$$m_{ij} = \begin{cases} 1 & \text{if arc } v_i v_j \in A \\ 0 & \text{otherwise} \end{cases}$$

Theorem 128 (Properties)

- *M is not necessarily symmetric*
- *The sum of any column of M is equal to the number of arcs directed towards v_j*
- *The sum of the entries in row i is equal to the number of arcs directed away from vertex v_i*
- *The $(i, j)-$entry of M^n is equal to the number of walks of length n from vertex v_i to v_j*

Definition 129 (Multiplicity of a pair)

The **multiplicity** of a pair x, y is the number $m_G^+(x, y)$ of arcs with initial endpoint x and terminal endpoint y. Let

$$m_G^-(x, y) = m_G^+(y, x)$$
$$m_G(x, y) = m_G^+(x, y) + m_G^-(x, y)$$

If $x \neq y$, then $m_G(x, y)$ is number of arcs with both x and y as endpoints. If $x = y$, then $m_G(x, y)$ equals twice the number of loops attached to vertex x. If $A, B \subset V$, $A \neq B$, let

$$m_G^+(A, B) = \{u : u \in U, u = (x, y), x \in A, y \in B\}$$
$$m_G(A, B) = m_G^+(A, B) + m_G^+(A, B)$$

Adjacency matrix of a multigraph

Definition 130 (Matrix associated with G)

If G has vertices $x_1, x_2, \dots, x_n$, then the **matrix associated** with G is

$$a_{ij} = m_G^+(x_i, x_j)$$

Definition 131 (Adjacency matrix)

The matrix $a_{ij} + a_{ji}$ is the **adjacency matrix** associated with G

Adjacency matrix (multigraph case)

Definition 132 (Adjacency matrix of a multigraph)

G an ℓ-graph, then the adjacency matrix $M_A = [m_{ij}]$ is defined as follows

$$m_{ij} = \begin{cases} k & \text{if arc there are } k \text{ arcs } (i,j) \in U \\ 0 & \text{otherwise} \end{cases}$$

with $k \leq \ell$

G undirected $\implies M_A(G)$ symmetric

$M_A(G)$ has nonzero diagonal entries if G is not simple.

Weighted adjacency matrices

Sometimes, adjacency matrices (typically for 1-graphs) have real entries, usually positive

This means that the arcs/edges have been given a weight

Theorem 133 (Number of walks of length n)

Let A be the adjacency matrix of a graph $G = (V(G), E(G))$, where $V(G) = \{v_1, v_2, \ldots, v_p\}$. Then the $(i, j)-$entry of A^n, $n \geq 1$, is the number of different walks linking v_i to v_j of length n in G.

(two walks of the same length are equal if their edges occur in exactly the same order)

Example: let A be the adjacency matrix of a graph $G = (V(G), E(G))$.

- the $(i, i)-$entry of A^2 is equal to the degree of v_i.
- the $(i, i)-$entry of A^3 is equal to twice the number of C_3 containing v_i.

Matrices associated to a graph/digraph

- Adjacency matrices
- Other matrices associated to a graph/digraph
- Linking graphs and linear algebra

Incidence matrix (undirected case)

Let $G = (V, E)$ be a graph of order p, and size q, with vertices $v_1, \ldots, v_p$, and edges $e_1, \ldots, e_q$

Definition 134 (Incidence matrix)

The incidence matrix is

$$B = B(G) = [b_{ij}]$$

is that $p \times q$ matrix in which

$$b_{ij} = \begin{cases} 1 & \text{if } v_i \text{ is incident with } e_j \\ 0 & \text{otherwise} \end{cases}$$

Theorem 135 (Incidence matrix and degrees)

The sum of the entries in row i of the incidence matrix is the degree of v_i in the graph

Incidence matrix (directed case)

Let $G = (V, A)$ be a digraph of order p and size q, with vertices $v_1, \ldots, v_p$ and arcs $a_1, \ldots, a_q$

Definition 136 (Incidence matrix)

The **incidence matrix** $B = B(G) = [b_{ij}]$ is a $p \times q$ matrix in which

$$b_{ij} = \begin{cases} 1 & \text{if arc } a_j \text{ is directed away from a vertex } v_i \\ -1 & \text{if arc } a_j \text{ is directed towards a vertex } v_i \\ 0 & \text{otherwise} \end{cases}$$

Spectrum of a graph

We will come back to this later, but for now..

Definition 137 (Spectrum of a graph)

The **spectrum** of a graph G is the spectrum (set of eigenvalues) of its associated adjacency matrix $M(G)$

This is regardless of the type of adjacency matrix or graph

Degree matrix

Definition 138 (Degree matrix)

The **degree** matrix $D = [d_{ij}]$ for G is a $n \times n$ diagonal matrix defined as

$$d_{ij} = \begin{cases} d_G(v_i) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

In an undirected graph, this means that each loop increases the degree of a vertex by two

In a directed graph, the term “degree” may refer either to indegree (the number of incoming edges at each vertex) or outdegree (the number of outgoing edges at each vertex)

Laplacian matrix

Definition 139 (Laplacian matrix)

$G = (V, A)$ a simple graph with n vertices. The **Laplacian** matrix is

$$L = D(G) - M(G)$$

where $D(G)$ is the degree matrix and $M(G)$ is the adjacency matrix

Laplacian matrix (continued)

G simple graph $\implies$ $M(G)$ only contains 1 or 0 and its diagonal elements are all 0

For directed graphs, either the indegree or outdegree is used, depending on the application

Elements of L are given by

$$\ell_{ij} = \begin{cases} d_G(v_i) & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise} \end{cases}$$

Distance matrix

Let G be a graph of order p with vertices $v_1, \ldots, v_p$

Definition 140 (Distance matrix)

The distance matrix $\Delta(G) = [d_{ij}]$ is a $p \times p$ matrix in which

$$\delta_{ij} = d_G(v_i, v_j)$$

Note $\delta_{ii} = 0$ for $i = 1, \ldots, p$

Property 141

- *M is not necessarily symmetric*
- *The sum of any column of M is equal to the number of arcs directed towards v_j*
- *The sum of the entries in row i is equal to the number of arcs directed away from vertex v_i*
- *The $(i,j)-$entry of M^n is equal to the number of walks of length n from vertex v_i to v_j*

Matrices associated to a graph/digraph

- Adjacency matrices
- Other matrices associated to a graph/digraph
- **Linking graphs and linear algebra**

Counting paths

Theorem 142

G a digraph and $M_A(G)$ its adjacency matrix. Denote $P = [p_{ij}]$ the matrix $P = M_A^k$. Then p_{ij} is the number of distinct paths of length k from i to j in G

Definition 143 (Irreducible matrix)

A matrix $A \in \mathcal{M}_n$ is **reducible** if $\exists P \in \mathcal{M}_n$, permutation matrix, s.t. $P^T AP$ can be written in block triangular form. If no such P exists, A is **irreducible**

Theorem 144

A irreducible $\iff$ $G(A)$ strongly connected

Theorem 145

Let A be the adjacency matrix of a graph G on p vertices. A graph G on p vertices is connected $\iff$

$$I + A + A^2 + \cdots + A^{p-1} = C$$

has no zero entries

Theorem 146

Let M be the adjacency matrix of a digraph D on p vertices. A digraph D on p vertices is strongly connected $\iff$

$$I + M + M^2 + \cdots + M^{p-1} = C$$

has no zero entries

Nonnegative matrix

$A = [a_{ij}] \in \mathcal{M}_n(\mathbb{R})$ **nonnegative** if $a_{ij} \geq 0 \; \forall i, j = 1, \ldots, n$; $\boldsymbol{v} \in \mathbb{R}^n$ nonnegative if $v_i \geq 0 \; \forall i = 1, \ldots, n$. **Spectral radius** of A

$$\rho(A) = \max_{\lambda \in \mathrm{Sp}(A)} \{|\lambda|\}$$

$\mathrm{Sp}(A)$ the **spectrum** of A

Perron-Frobenius (PF) theorem

Theorem 147 (PF – Nonnegative case)

$0 \leq A \in \mathcal{M}_n(\mathbb{R})$. Then $\exists \mathbf{v} \geq \mathbf{0}$ s.t.

$$A\mathbf{v} = \rho(A)\mathbf{v}$$

Theorem 148 (PF – Irreducible case)

Let $0 \leq A \in \mathcal{M}_n(\mathbb{R})$ irreducible. Then $\exists \mathbf{v} > \mathbf{0}$ s.t.

$$A\mathbf{v} = \rho(A)\mathbf{v}$$

$\rho(A) > 0$ and with algebraic multiplicity 1. No nonnegative eigenvector is associated to any other eigenvalue of A

Primitive matrices

Definition 149

$0 \leq A \in \mathcal{M}_n(\mathbb{R})$ **primitive** (with **primitivity index** $k \in \mathbb{N}_+^*$) if $\exists k \in \mathbb{N}_+^*$ s.t.

$$A^k > 0,$$

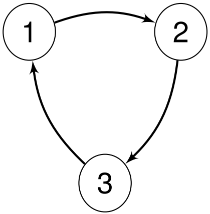
with k the smallest integer for which this is true. A **imprimitive** if it is not primitive

A primitive $\implies$ A irreducible; the converse is false

Theorem 150

$A \in \mathcal{M}_n(\mathbb{R})$ irreducible and $\exists i = 1, \ldots, n$ s.t. $a_{ii} > 0 \implies A$ primitive

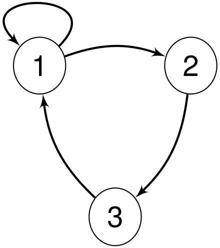
Here d is the index of imprimitivity (i.e., the number of eigenvalues that have the same modulus as $\lambda_p = \rho(A)$). If $d = 1$, then A is primitive. We have that $d = \text{gcd}$ of all the lengths of closed walks in $G(A)$



Adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Closed walks in $G(A)$ (lengths): $1 \to 1$ (3), $2 \to 2$ (3), $2 \to 2$ (3) $\implies \gcd = 3 \implies d = 3$ (here, all eigenvalues have modulus 1)



$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Closed walk $1 \to 1$ has length 1 $\implies$ gcd of lengths of closed walks is 1 $\implies$ A primitive

Let $\mathbf{0} \leq A \in \mathcal{M}_n$

Theorem 151

A primitive $\implies \exists\, 0 < k \leq (n-1)n^n$ *such that* $A^k > 0$

Theorem 152

If A is primtive and the shortest simple directed cycle in $G(A)$ has length s, then the primitivity index is $\leq n + s(n-1)$

Theorem 153

A primitive $\iff A^{n^2-2n+2} > \mathbf{0}$

Theorem 154

If A is irreducible and has d positive entries on the diagonal, then the primitivity index $\leq 2n - d - 1$

Theorem 155

$\mathbf{0} \leq A \in \mathcal{M}_n$, $\lambda_P = \rho(A)$ *the Perron root of* A, $\boldsymbol{v}_P$ *and* $\boldsymbol{w}_P$ *the corresponding right and left Perron vectors of* A, *respectively,* d *the index of imprimitivity of* A *(with* $d = 1$ *when* A *is primitive) and* $\lambda_j \in \sigma(A)$ *the spectrum of* A, *with* $j = 2, \ldots, n$ *unless otherwise specified (assuming* $\lambda_1 = \lambda_P$*)*