Basic notions, notation and convention

Graph $G = G(V, E)$

V: vertex/node set, $n = |V|$

E: edge/link set, $m = |E|$, $E \subset V \times V$

G is simple if there is at most one edge between any two vertices.

A k-partite, $k > 1$, is a graph $G\left(V = \bigcup_{j=1:k} V_j,\ E = \bigcup_{i \neq j} E_{ij}\right)$ where $E_{ij} \subset V_i \times V_j$ and $V_i \cap V_j = \emptyset$ if $i \neq j$. By the edge set, each vertex subset V_i is an independent set. In particular, a bipartite is a graph $G(V_1, V_2, E_{1,2})$ with $E_{12} \subset V_1 \times V_2$.

A subgraph induced by a vertex subset U is $G(U, E \cap (U \times U))$. A bipartite subgraph induced by two non-overlapping vertex sets V_1 and V_2 is $G(V_1, V_2, E \cap (V_1, V_2))$.

Neighborhoods and neighbor graphs of vertex node $v \in V$:

$\mathcal{N}(v) = \{u : (u, v) \in E, u \neq v\}$, exclusion of v

$\mathcal{N}[v] = \mathcal{N}(v) \cup \{v\}$, inclusion of v

$G(v)$ denotes the subgraph induced by $\mathcal{N}(v)$; $G[v]$, by $\mathcal{N}[v]$.

Two basic vertex functions or neighborhood feature descriptors:

* $d(v) = |\mathcal{N}(v)|$ is the degree of v is the number of
* $\text{lcc}(v)$ is the *local cluster coefficient* at vertex v with $d(v) > 1$,

$$\text{lcc}(v) = \frac{|E(G(v))|}{\binom{d(v)}{2}} \leq 1, \quad d(v) > 1, \quad v \in V. \tag{1}$$

That is, it is is defined as the edge density of the neighbor graph $G(v)$. The LCC concept is introduced by D. J. Watts and S. Strogatz in 1998.

The $n \times n$ identify matrix is I, with $e_j = I(:, j)$. The constant-1 vector is $e = \sum_{j=1:n} e_j$

Often, graph operations or relations can be described clearly via adjacency matrices. Adjacency matrix A of graph G: $A \iff G$

- Use a particular vertex-index mapping: $V \to \{1, 2, \cdots, n\}$
- $A(i, j) \neq 0 \iff (i, j) \in E$, $A(i, i) \neq 0 \iff$ vertex i has a self-loop.

- Each row/column corresponds to a vertex
  - $Ae_j = A(:, j)$: in column-j, the nonzero elements represent the outgoing edges (and the neighbor nodes) from vertex j
  - $e_i^{\mathrm{T}} A = A(i, :)$: in row-$i$, the nonzero elements represent the incoming edges (and the neighbor nodes) from vertex i

Graph G is also uniquely specified by its incidence matrix $B_{n\times m}$,

$$B(:, \ell) = e_i - e_j, \quad \ell = (i, j) \in E. \tag{2}$$

Hadamard multiplication and division between two arrays of the same dimensions are elementwise operations and denoted as $\otimes$ (or $.\times$) and ./ respectively.

Analysis

Let $G(V, E)$ be a simple, undirected, unweighted, connected graph. Assume A is the adjacency matrix of G by some vertex-to-index mapping. Then, any subset U of V identifies with a subset of $\{1, \cdots, n\}$.

1. [T/F/M] $A^{\mathrm{T}} = A$.

   [T/F/M] The subgraph induced by $U \subset V$ is represented by $A(U, U)$.

   [T/F/M] The bipartite subgraph induced by $V_1, V_2 \subset V$, $V_1 \cap V_2 = \emptyset$, is represented by $A(V_1, V_2)$.

   [T/F/M] without self-loops, $m = \mathrm{nnz}(A)/2$.

   [T/F/M] without self-loops, the number of vertices with odd degrees is even. `(The handshaking lemma)`

2. Neighborhood.

   [T/F/M] At any vertex v, $\mathcal{N}[v]$ is not an independent set, whereas $\mathcal{N}(v)$ may be an independent set.

   [T/F/M] If $G[v]$ is a clique, so is $G(v)$, and vice versa.

3. Triangles incident at vertex v and neighborhood graph $G(v)$.

   [T/F/M] Every edge between two neighbors of v is the base of a triangle incident at v.

   [T/F/M] Denote by $\#C_3(v)$ the total number of triangles incident at node v. Then,

   $$\#C_3(v) = |E(G(v))| \leq \binom{d(v)}{2}.$$

   By the equality, the LCC coefficient is the ratio of the existing number of triangles incident at v to the number of all potentially possible triangles incident at v.

   Optional. An Mycielski graph is triangle-free by construction. It has the largest edge set size among triangle-free graphs of the same size.

   Optional. Find or construct at least three more (types of) triangle-free graphs, not including star graphs or Mycielski graphs.

4. Degree expression and LCC expressions via matrix-vector operations,

   [T/F/M]
   $$d = d(1:n) = A\,e \tag{3}$$

   [T/F/M]
   $$\mathrm{lcc}(1:n) = 2[\mathrm{diag}(A^3)e]\,./(d \odot (d-1)) \tag{4}$$

   [T/F/M]
   $$\mathrm{lcc}(1:n) = [A^2 \odot A)e]\,./\,(d \odot (d-1)). \tag{5}$$

5. [T/F/M] Connectivity and reachability. Given that G is connected, for any pair of vertices u and v, there exists an integer k, $k \leq \mathrm{diam}(G)$, such that $R_k(i,j) = [\sum_{j=1:k} A^j](u,v) > 0$.

6. Determine the length of the shortest path(s) between any pair of nodes $u, v \in V$ by the sequence $\{A^k(u,v), k = 1, 2, \cdots, n-1\}$.

7. Verify the following relationship between the degree vector, the adjacency matrix and the incidence matrix, equalities,

$$BB^{\mathrm{T}} = \mathrm{diag}(d) - A, \quad d = A\,e \tag{6}$$

The gram (product) matrix BB^{T} is actually the Laplacian matrix of graph G and denoted as L.

8. [Optional to undergrads.] Verify that the following quadratic function on a connected graph G is nonnegative,

$$x^{\mathrm{T}} L x \geq 0. \tag{7}$$

The equality holds if and only if x is a constant vector.