Experiments

1. Objective: data gathering and initial inspection.

   ◦ gather at least 3 connected real-world networks, each with at least 1,000 nodes.

     Use a provided name or give a name to each network, describe briefly the semantic meaning of the vertices and edges for each network $G(V, E)$, and cite the data sources.

   ◦ format conversion from the incidence matrix to the adjacency matrix,

   ◦ make a summary table of the networks with the basic feature elements $n = |V|$, directed/undirected, weighted/unweighted, $m = |V|$, $\max(d_{\text{in}})$, $\min(d_{\text{min}})$, $\text{avg}(d)$, degree-distribution type, and diameter (if available or inexpensive to compute).

   For each network,

   ◦ plot a histogram for the degree distribution,

   ◦ plot a histogram for the lcc distribution.

   (∗) One can use or modify the provided MATLAB codes, or other available and reliable codes, for computing the degrees and the local cluster coefficients and displaying their PDF histograms. Each histogram has at least 10 bins.

2. Objective: construction of a related network.

   Let $G_1 = G(V, E)$ be an undirected, unweighted graph at input. Let $G_k = G(V, E_k)$ be the constructed graph at output with an integer parameter k, $k > 1$,

   $$(u, v) \in E_k \iff (v_j, v_{j-1}) \in E, \qquad 1 \le j \le k, \qquad v_0 = v, \quad v_k = u,$$

   where no vertex is visited more than once. In other words, G_k describes the k-step connectivity.

   ◦ For each real-world network G_1 gathered for the preceding problem, generate G_2 (hint: get the adjacency matrix A_2 from A_1),

   ◦ Show the difference of G_2 from G_1 in degree distribution and LCC distribution.

3. Objective: get familiar with 5 types of synthetic graphs/networks:

   (a) Erd̈Os and Rényi (ER) graphs,
   (b) Watts and Strogatz (WS), a.k.a., small-world graphs
   (c) Barabási and Albert (BA), a.k.a., scale-free graphs
   (d) Geometric random (GR) graphs
   (e) Topologically determined graphs (vs random or stochastic)

   To do:

◦ make a summary table of the datasets provided by the function `select_graph_matrix_A`.

◦ plot the degree and LCC distributions and comment on distinct signatures

4. Objective: graph representation of feature data points/vectors in a metric space. Let k be a positive integer parameter.

   - gather at least three sets of feature data vectors
   - construct k-neighbor graphs for each, assuming the conventional Euclidean distance (the out-degree is constant k)
   - plot the in-degree distribution
   - [Optional:] propose (and name) a local density measure similar to LCC for undirected graphs, and plot a histogram of the proposed density measure.

5. Objective: get familiar with the use of graph Laplacian spectrum.

   Assume graph G is undirected and connected.

   To experiment:

   - ▷ Graph cut by the Fiedler vector: use the demo script `demo_Fiedler_pair/m` on two datasets and make brief inspection comments.
   - ▷ Graph embedding in a spectral subspace: use the demo script `demo_Laplacian_embedding.m` for low-dimensional embedding and make brief inspection comments.
   - ▷ [Optional to undergrads.] Apply the Fiedler cut to a real-world network.
   - ▷ [Optional.] Apply the Laplacian embedding to a real-world network.