Analysis

1. (a) True.
   For an undirected graph, the adjacency matrix is symmetric, meaning that $A^T = A$.

   (b) True.
   The subgraph induced by a subset $U \subset V$ is represented by the submatrix $A(U, U)$, which corresponds to the vertices in U and the edges between them.

   (c) False.
   $A(V_1, V_2)$ describes only the relationship between the vertices in V_1 and V_2, i.e., the edges between these two sets. However, it doesn't capture the absence of internal edges within V_1 or V_2. To properly represent a bipartite subgraph, we need to construct the matrix $A(\{V_1, V_2\}, \{V_1, V_2\})$, where the block diagonal elements (representing internal edges within V_1 and V_2) are zero, ensuring no internal edges exist within these sets.

   (d) True.
   The number of non-zero elements in A, denoted as $\text{nnz}(A)$, counts each edge twice in an undirected graph (once for each direction), so the number of edges m is half of the non-zero elements in A.

   (e) True.
   According to the handshaking lemma, in any undirected graph, the sum of vertex degrees is twice the number of edges, and the number of vertices with odd degrees must be even.

2. Neighborhood.

   (a) Maybe.
   The statement says that $\mathcal{N}[v]$ is not an independent set for any vertex v. However, there are exceptions: in a disconnected graph, a vertex v_0 might be isolated, making $\mathcal{N}[v_0] = \{v_0\}$, which is trivially an independent set. If we modify the statement to say "...if $\mathcal{N}[v]$ is not an independent set..." or similar, it would become true. This is because, in the general case, $\mathcal{N}[v]$ includes edges between v and its neighbors, and hence cannot be an independent set.

   (b) True.
   If $G[v]$, which includes the vertex v and its neighborhood, forms a clique, this implies that all vertices in $G[v]$, including v, are pairwise adjacent. Since $G(v)$ only excludes v but includes all its neighbors, $G(v)$ will also be a clique. Similarly, if $G(v)$ is a clique, adding v back into the graph will form a clique in $G[v]$.

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

   (a) True.
   Every edge between two neighbors of v forms the base of a triangle that is incident at v. This is because for a triangle to exist, the two neighbors must be connected to each other, and they are both adjacent to v, forming a triangle.

(b) True.
The number of triangles incident at node v, denoted as $\#C_3(v)$, is equal to the number of edges in $G(v)$, the neighborhood graph of v. It is bounded by the maximum possible number of triangles, which is $\binom{d(v)}{2}$, where $d(v)$ is the degree of v. This is the maximum number of edges that can exist between the neighbors of v, and thus the upper bound on the number of triangles.

(c) Optional: False.
While a Mycielski graph is triangle-free by construction, it does not have the largest edge set size among triangle-free graphs of the same size. For example, the Turan graph $T(n = 11, m = 30)$ has more edges than the Mycielski graph M_5 (with $n = 11, m = 20$), while both are triangle-free. The Turan graph maximizes the number of edges for triangle-free graphs.

(d) Optional:
Three other types of triangle-free graphs (not including star graphs or Mycielski graphs) could include:

i. Bipartite graphs: These graphs are triangle-free because they can be split into two disjoint sets, with no edges within each set, only between them.
ii. Cycle graphs with more than three vertices (e.g., C_4, C_5, etc.)**: Any cycle graph C_n where $n \geq 4$ does not contain any triangles.
iii. Grid graphs: A 2D grid graph does not contain any triangles because each node is connected only to its vertical and horizontal neighbors.

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

(a) True.
The degree vector d for a graph can be expressed as $d = Ae$, where A is the adjacency matrix and e is the all-ones vector. This is because each entry d_i in the degree vector is the sum of the entries in the i-th row of A, which corresponds to the number of neighbors for vertex i.

(b) False.
The expression $\text{lcc}(1:n) = 2[\text{diag}(A^3)e]\ ./\ (d \otimes (d-1))$ is incorrect because the factor of 2 is unnecessary. The term $\text{diag}(A^3)$ counts the number of 3-step walks that return to the starting vertex, but this counts each triangle twice (once for each of the three vertices). Similarly, $d \otimes (d-1)$ also counts combinations of neighbors twice. These factors of 2 cancel out, so the correct expression should be:

$$\text{lcc}(1:n) = [\text{diag}(A^3)e]\ ./\ (d \otimes (d-1)).$$

(c) True.
The expression $\text{lcc}(1:n) = [A^2 \otimes A]e\ ./\ (d \otimes (d-1))$ is correct. Here, $A^2 \otimes A$ captures the overlap between 2-step and 1-step walks, which counts triangles twice (once for each of the two other vertices involved in the triangle). Dividing by $d \otimes (d-1)$ normalizes by the maximum possible number of triangles, correctly calculating the local clustering coefficient (LCC).

5. Connectivity and reachability.

   True.
   The statement is true because in a connected graph G, there must be a path between any two vertices u and v. The diameter of the graph $\text{diam}(G)$ is the maximum distance between any pair of vertices in the graph. For any pair of vertices u and v, there must exist a shortest path between them of some length k, where $k \leq \text{diam}(G)$. The sum of powers of the adjacency matrix A (i.e., $\sum_{j=1}^{k} A^j$) counts all walks of length up to k. Since u and v are connected, the entry (u, v) in this matrix sum will be greater than 0, indicating the existence of a path.

6. Answer: The length of the shortest path between nodes u and v is the smallest integer k, such that $A^k(u, v) > 0$.

   - $A^k(u, v)$ represents the number of walks of length k from u to v.
   - The shortest path is determined by finding the smallest k for which $A^k(u, v) > 0$, indicating that a path of length k exists between u and v.
   - This works because for a connected graph, there is always a path between u and v within $k \leq n - 1$, where n is the number of nodes in the graph.

7. Answer: To verify the given relationship, we need to confirm that the Gram matrix $BB^\top$, formed by the incidence matrix B, is equivalent to the Laplacian matrix of the graph, expressed as $\text{diag}(d) - A$, where $d = Ae$ and A is the adjacency matrix.

   We begin by analyzing the case of an undirected graph.

   (a) **Incidence Matrix B:** The incidence matrix B is defined as follows:
      - $B(:, l) = e_i - e_j$ if edge $l = (i, j) \in E$, where e_i and e_j are standard basis vectors.
      - For each edge l connecting vertices i and j, column l of B has 1 at row i and -1 at row j, with zeros elsewhere.

   (b) **Product $BB^\top$:** The product $BB^\top$, also called the Gram matrix, has the following properties:
      - The diagonal entries $(BB^\top)_{ii}$ give the degree of vertex i, i.e., $(BB^\top)_{ii} = \deg(i)$.
      - The off-diagonal entries $(BB^\top)_{ij}$ are -1 if vertices i and j are connected by an edge, and 0 otherwise.

      Thus, $BB^\top$ encodes degree information on the diagonal and adjacency information (with negative signs) off-diagonal.

   (c) **Laplacian Matrix L:** The Laplacian matrix L is defined as:

   $$L = \text{diag}(d) - A$$

   where:

- diag(d) is the diagonal matrix of vertex degrees.
- A is the adjacency matrix.

From the structure of $BB^\top$, it is evident that it matches the Laplacian matrix structure:

- The diagonal elements represent vertex degrees.
- The off-diagonal elements are -1 for adjacent vertices.

(d) **Conclusion for Undirected Graphs:** We conclude that:

$$BB^\top = \text{diag}(d) - A$$

Therefore, the Gram matrix $BB^\top$, formed by the incidence matrix B, is indeed the Laplacian matrix for undirected graphs.

Now, consider the case of directed graphs.

For directed graphs, the adjacency matrix A is asymmetric, meaning that an edge from vertex i to vertex j does not imply the presence of an edge from j to i. The Laplacian matrix for directed graphs is typically defined as:

$$L = D_{\text{out}} - A$$

where D_{out} is the out-degree matrix and A is the asymmetric adjacency matrix.

However, the product $BB^\top$ is always symmetric, because:

$$(BB^\top)^\top = BB^\top$$

Thus, the Gram matrix $BB^\top$ cannot represent the Laplacian matrix for directed graphs, as the Laplacian for directed graphs must be asymmetric to reflect the directionality of edges. The symmetry of $BB^\top$ contradicts the asymmetry of the adjacency matrix A and the directed Laplacian.

Final Conclusion:

- For undirected graphs, the relationship $L = BB^\top$ holds, as both the Laplacian matrix and $BB^\top$ are symmetric and accurately reflect the graph's structure.
- For directed graphs, the relationship $L = BB^\top$ does **not** hold because $BB^\top$ is symmetric, whereas the directed graph Laplacian is asymmetric due to the structure of the directed adjacency matrix.

8. [Optional to undergrads.]

Answer: To verify this, we use the properties of the Laplacian matrix L, which is defined as $L = \text{diag}(d) - A$, where $\text{diag}(d)$ is the diagonal matrix of vertex degrees, and A is the adjacency matrix of the graph G.

Step 1: Expand the quadratic form

The quadratic form $x^\top Lx$ can be written as:

$$x^\top Lx = \sum_{i=1}^{n} \sum_{j=1}^{n} L_{ij} x_i x_j$$

where L_{ij} represents the entries of the Laplacian matrix L.

Step 2: Substituting the definition of the Laplacian

Recall that the Laplacian matrix L is defined as $L = \text{diag}(d) - A$, where:

$$L_{ij} = \begin{cases} d_i, & \text{if } i = j, \\ -1, & \text{if } (i,j) \in E \text{ (i.e., edge exists from vertex } i \text{ to vertex } j), \\ 0, & \text{otherwise.} \end{cases}$$

Thus, the quadratic form becomes:

$$x^\top Lx = \sum_{i=1}^{n} d_i x_i^2 - \sum_{(i,j)\in E} x_i x_j$$

Here's how this works:

- The first sum $\sum_{i=1}^{n} d_i x_i^2$ accounts for the degree of each vertex i, where d_i is the degree of vertex i. This gives the total contribution from each vertex's own value x_i^2.
- The second sum $\sum_{(i,j)\in E} x_i x_j$ represents the adjacency relationships between vertices. For each edge $(i,j) \in E$, we subtract the product $x_i x_j$, reflecting the interaction between adjacent vertices.

Step 3: Rewriting the quadratic form for directed graphs

For **directed graphs**, the expression $x^\top Lx$ does not directly rewrite into a perfect sum of squared differences $(x_i - x_j)^2$ because $(i,j) \in E$ does not imply $(j,i) \in E$. However, we can still show that:

$$x^\top Lx \geq \frac{1}{2} \sum_{(i,j)\in E} (x_i - x_j)^2$$

where equality holds **only** when for every $(i,j) \in E$, the reverse edge $(j,i) \in E$ also exists, i.e., the graph is effectively bidirectional for those edges.

Here's why: for directed graphs, if $(i,j) \in E$ but $(j,i) \notin E$, the missing term $-x_i x_j$ from the reverse edge would otherwise reduce the quadratic form. Since this missing term would be negative (i.e., $-x_i x_j$ is negative), the overall sum in $x^\top Lx$ is larger when the reverse edge is absent, thus resulting in the inequality.

In contrast, for **undirected graphs**, where $(i,j) \in E$ implies $(j,i) \in E$, the sum can be rewritten as:

$$x^\top Lx = \frac{1}{2} \sum_{(i,j)\in E} (x_i - x_j)^2$$

with no inequality.

Step 4: Nonnegativity of $x^\top Lx$

Since $(x_i - x_j)^2 \geq 0$ for all pairs of vertices i and j, it follows that:

$$x^\top Lx \geq \frac{1}{2} \sum_{(i,j)\in E} (x_i - x_j)^2 \geq 0.$$

The quadratic form is nonnegative because it is bounded below by the sum of squared differences, and that sum is itself nonnegative.

Step 5: Equality Case

The equality $x^\top Lx = 0$ holds if and only if:

$$(x_i - x_j)^2 = 0 \quad \forall (i,j) \in E.$$

This implies that $x_i = x_j$ for all connected vertices i and j.

For undirected graphs: This means $x_i = x_j$ for all i and j, since each edge (i,j) has the reverse edge (j,i) due to the undirected nature of the graph. Hence, x must be of the form:

$$x = c \cdot \mathbf{1}$$

where c is some constant scalar, and $\mathbf{1}$ is the all-ones vector.

For directed graphs: The equality $x^\top Lx = 0$ holds if $x_i = x_j$ for all directed edges $(i,j) \in E$. However, in cases where $(i,j) \in E$ but $(j,i) \notin E$, the missing reverse term $-x_i x_j$ implies that the sum $x^\top Lx$ is strictly larger than $\frac{1}{2}\sum(x_i - x_j)^2$. Hence, the equality holds only when $(i,j) \in E$ and $(j,i) \in E$ for all such edges.

Conclusion

We have shown that $x^\top Lx \geq \frac{1}{2}\sum(x_i - x_j)^2 \geq 0$, with equality if and only if all edges $(i,j) \in E$ have their reverse edges $(j,i) \in E$, or equivalently, the graph is effectively undirected for those edges. In this case, for both directed and undirected graphs, the quadratic form achieves equality only when $x = c \cdot \mathbf{1}$, i.e., x is a constant vector.

References