Important Properties and Applications of the Adjacency Matrix (20250310)

Shuliang Bai

1 Section 1. Counting Walks with Adjacency Matrix Powers

Theorem 1. *Let $G = (V, E)$ be a simple graph with n vertices, and let A be its $n \times n$ adjacency matrix. Then for any two vertices i and j, the entry $(A^k)_{ij}$ counts the number of walks of length k from vertex i to vertex j.*

Walk: A sequence of vertices where each consecutive pair is connected by an edge. Vertices and edges can be repeated.

Path: A walk where no vertex is repeated (except possibly the first and last vertex if it's a cycle).

The Adjacency Matrix Power Theorem

$$(A^k)_{ij}$$

counts walks — not paths. This is because matrix multiplication accumulates all sequences of edges between two vertices, including those that revisit vertices and edges. It does not "remember" whether a vertex has already been visited.

Proof. We prove by induction on k.
Base case: For $k = 1$, $(A)_{ij} = 1$ if and only if there is an edge from i to j, which is exactly the number of walks of length 1. $(A)_{ii} = 0$ as there is no walk of length 1 from a vertex to itself.
Inductive step: Assume the claim holds for k. To count walks of length $k + 1$ from i to j, we can extend any walk of length k ending at vertex ℓ by adding an edge from ℓ to j. Thus:

$$(A^{k+1})_{ij} = \sum_{\ell=1}^{n} (A^k)_{i\ell} A_{\ell j}$$

This is the standard matrix product rule, so

$$A^{k+1} = A^k \cdot A$$

The result follows by induction. □

Theorem 2 (Matrix Criterion for Connectivity)**.** *Let G be a simple graph with n vertices, and let A denote its adjacency matrix. Then G is connected* if and only if *the matrix power $(I + A)^{n-1}$ has all entries strictly positive.*

Proof. *Proof Sketch.* The connectivity criterion can be understood through these key observations:

1. **Path Length Bound**: Any path in an n-vertex graph contains at most $n-1$ edges. If no path exists between two vertices within $n-1$ edges, the graph is disconnected.

2. **Why A^{n-1} Fails**: While A^{n-1} counts paths of *exactly* $n-1$ edges, it misses shorter paths. This insufficiency motivates the need for $(I+A)^{n-1}$.

3. **The Role of I**: Define the modified adjacency matrix $A' = I + A$, where the identity matrix I adds self-loops ($A'_{i,i} = 1$). This allows:

   - Staying at the same vertex during path traversal
   - Counting paths of length *at most* k in $(A')^k$

4. **Final Criterion**: $(I+A)^{n-1}$ aggregates all paths of length $\leq n-1$. If any entry $(A')^{n-1}_{i,j} = 0$, vertices i and j are disconnected.

Remark: This mirrors the Bellman-Ford algorithm's use of the $n-1$ edge bound for detecting negative cycles. □

Applications

- **Network Analysis:** Counting walks of different lengths to study how well-connected two nodes are in social networks.
- **Transport Systems:** Estimating the number of alternative routes between locations in a transport network.
- **Bioinformatics:** Analyzing pathways in protein interaction networks to identify potential biological processes.

2 Section 2. Properties of Eigenvalues

Let G be a graph with adjacency matrix A, and let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be its eigenvalues. Key properties and applications are outlined below. First, summing over all vertices, we obtain:

$$\sum_i \left(A^k\right)_{ii} = \text{ total number of closed walks of length } k \text{ in } G$$

Second, by leveraging the spectral decomposition of A, we can express this sum in terms of the eigenvalues of A, which provides a direct bridge between graph structure and spectral properties.

Theorem 3. *Closed Walk Counting*

In simpler terms, a closed walk starts at some vertex, follows edges, and eventually returns to the starting vertex.

For any integer $k \geq 0$,

$$\sum_{i=1}^{n} \lambda_i^k = \textit{Number of closed walks of length } k.$$

- *Case* $k = 0$: Trivially equals n, the number of vertices.
  Application: Verifies graph size in data structures.
- *Case* $k = 1$: Sum of eigenvalues equals zero.
  Application: Validates adjacency matrix construction in network simulations.
- *Case* $k = 2$: **Edge Detection via Quadratic Moment**:

  $$\frac{1}{2}\sum_{i=1}^{n} \lambda_i^2 = \text{Number of edges in } G.$$

  This formula shows how the global connectivity (number of edges) is encoded in the eigenvalues of the adjacency matrix. In some sense, the eigenvalues "know" how much total connection (edge count) exists in the graph. **Corollary**: A graph is edgeless iff all eigenvalues are zero.
  Application: Efficiently detect sparse/dense graphs in computational biology.
- *Case* $k = 3$: **Triangle Counting via Cubic Moment**:

  $$\frac{1}{6}\sum_{i=1}^{n} \lambda_i^3 = \text{Number of triangles in } G.$$

  Corollary: A graph is triangle-free iff this sum is zero.
  Application: Identify clustering coefficients in social network analysis.

2.1 k=2

Let A be the adjacency matrix of a graph G with n vertices. The number of edges in G is counted by $1/2 \sum A_{ij} = 1/2 \sum A_{ij}^2$. Thus, we need to prove that the sum of the squares of the eigenvalues of A equals the sum of the squares of all entries in A:

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i,j} A_{ij}^2$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A.

By defintion, the Frobenius norm of A is defined as:

$$\|A\|_F^2 = \sum_{i,j} A_{ij}^2$$

This expression just means that we add up the squares of all elements in the matrix.

For any symmetric matrix A, the Frobenius norm can also be computed using the eigenvalues:

$$\|A\|_F^2 = \sum_{i=1}^{n} \lambda_i^2$$

This is a standard result from linear algebra, based on the fact that A is orthogonally diagonalizable (since A is symmetric), i.e., $A = Q\Lambda Q^T$, where Λ is the diagonal matrix containing $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Proof. This result follows from the fact that for symmetric matrices, the matrix can be diagonalized. In other words, the matrix A can be written as:

$$A = Q\Lambda Q^T$$

Where: - Q is an orthogonal matrix (meaning $Q^TQ = I$, the identity matrix), - Λ is a diagonal matrix whose diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

So, when we compute the Frobenius norm of A, we get:

$$\|A\|_F^2 = \sum_{i,j} A_{ij}^2 = \text{Tr}\left(A^TA\right) = \text{Tr}\left(A^2\right)$$

Since $A^2 = Q\Lambda Q^TQ\Lambda Q^T = Q\Lambda^2Q^T$ (because $Q^TQ = I$, the identity matrix), the trace of A^2 is simply the sum of the diagonal elements of Λ^2, which are λ_i^2.

Therefore:

$$\sum_{i,j} A_{ij}^2 = \text{Tr}\left(A^2\right) = \sum_{i=1}^{n} \lambda_i^2$$

This completes the proof.

□

2.2 k=3, Application: Triangle Detection and Clustering Coefficients

There are two main types of clustering coefficients we look at:

Local Clustering Coefficient: This refers to the clustering around a single node. Think of it as checking, for each person in a group, how many of their friends know each other. We calculate it by counting the number of actual triangles (where three people are mutually connected) that a person is part of, and dividing that by the maximum number of triangles that could exist with that person's friends.

A triangle is the smallest possible closed structure involving 3 vertices.

For example, if you have three friends, and they all know each other, that's a perfect triangle, and your local clustering coefficient will be high. But if only two of your friends know each other, and the third doesn't, your clustering coefficient will be lower.

This "friend-of-a-friend" connection is a 3-way relationship, which is why triangles — not 4-cycles or longer loops — matter here.

Global Clustering Coefficient: This looks at the entire network or graph and measures the overall tendency of nodes to form triangles. It's like looking at the whole social circle and figuring out how many groups are tightly connected compared to how many possible connections there could have been in the network.

Clustering Coefficient Definition: The global clustering coefficient C measures the density of triangles in a network:

$$C = \frac{3 \times \text{Number of triangles}}{\text{Number of connected triples}}$$

Number of Triangles: The number of complete triangles (i.e., triplets of nodes that are mutually connected).

Number of Connected Triples: A connected triple is a set of three nodes where at least two nodes are connected by an edge.

If you have a high global clustering coefficient, it means the network has many tightly-knit groups where everyone knows each other. On the other hand, a low global clustering coefficient suggests that the network is more dispersed or 'loose,' with fewer connections between neighbors.

Why is this important? The clustering coefficient is important in many real-world scenarios. In social networks, it helps us understand how people form communities or groups of friends. In biological networks, like protein interaction networks, it can reveal how closely related different proteins or genes are. And in computer networks, it can show us how well connected the system is, or how quickly information might spread across the network.

In summary, the Clustering Coefficient helps us analyze the structure of a network—whether it's tight and interconnected or spread out and disconnected—and this can give us valuable insights into how networks operate in different fields."

The cubic moment of eigenvalues provides a powerful tool for analyzing social networks:

$$\frac{1}{6}\sum_{i=1}^{n} \lambda_i^3 = \text{Number of triangles in } G$$

Here, the cubic eigenvalue sum directly quantifies the numerator.

Why Eigenvalues?: For large networks (e.g., Facebook communities), explicitly counting triangles is computationally expensive ($O(n^3)$). Calculating $\sum \lambda_i^3$ via matrix diagonalization reduces complexity to $O(n^3)$ *theoretically*, but optimized algorithms (e.g., power iteration) achieve approximations in $O(n^2)$.

Algorithmic Advantage: Spectral methods avoid edge-by-edge traversal, making them suitable for:

- Real-time community detection in streaming networks
- Anomaly detection (sudden drops in $\sum \lambda_i^3$ may indicate broken social ties)

Limitations: Requires the graph to be undirected and unweighted. For weighted/directed graphs, modified adjacency matrices are needed.

Proof. $(A^k)_{ij}$ describes the number of walks of length k, thus, $(A^k)_{ii}$ is the number of closed walks for vertex i. The total number of closed walks for all vertices is then $\sum_i (A^k)_{ii} = \mathrm{Tr}(A^3) = \sum_{i=1}^n \lambda_i^3$. Each closed walk in the above count one triangle 6 times. $\square$

Bipartite Graphs: For bipartite graphs,

$$\sum_{i=1}^{n} \lambda_i^k = 0 \quad \text{for all odd } k.$$

Application: Classify bipartite structures in recommendation systems (e.g., user-item networks).

3 Estimate of Eigenvalues

Let us recall the definition of matrix norms.

- Let A^{H} be the conjugate transpose of the square matrix A , so that $(a_{ij})^{\mathrm{H}} = (\bar{a}_{ji})$, then the spectral norm is defined as the square root of the maximum eigenvalue of $A^H A$, i.e.,

$$\begin{aligned} \|A\|_2 &= \left(\text{ maximum eigenvalue of } A^H A \right)^{1/2} \\ &= \max_{|x|_2 \neq 0} \frac{|\ \mathrm{A}\mathbf{x}|_2}{|\mathbf{x}|_2} \end{aligned}$$

  This matrix norm is implemented as Norm[m, 2].

- The Frobenius norm, sometimes also called the Euclidean norm (a term unfortunately also used for the vector L^2-norm), is matrix norm of an $m \times n$ matrix A defined as the square root of the sum of the absolute squares of its elements,

$$\|\mathrm{A}\|_F \equiv \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$$

  (Golub and van Loan 1996, p. 55). The Frobenius norm can also be considered as a vector norm. It is also equal to the square root of the matrix trace of AA^H, where A^H is the conjugate transpose, i.e.,

$$\|\mathrm{A}\|_F = \sqrt{\mathrm{Tr}\left(\mathrm{AA}^{\mathrm{H}}\right)}$$

- The natural norm induced by the L1-norm is called the maximum absolute column sum norm a

$$\|\mathrm{A}\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

Let us look at the **spectral norm** of a matrix A, denoted $\|A\|_2$, is defined as:

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

This represents the maximum stretch factor by which A can transform a vector.

Definition 2.1. For M a self-adjoint operator over an inner product space $\mathcal{V} = \mathbb{R}^n$, and a nonzero $x \in \mathcal{V}$, let Rayleigh's quotient be

$$R(x) = \frac{\langle x, Mx \rangle}{\|x\|^2}$$

Theorem 2.1 (Courant-Fischer Theorem). Let M be a self-adjoint operator over an inner product space $\mathcal{V} = \mathbb{R}^n$ with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ repeated according to their multiplicities and corresponding eigenvectors $e_1, \ldots, e_n$. Then

$$\lambda_1 = \max_{x \neq 0} R(x)$$

and in general

$$\lambda_k = \max_{\substack{x \neq 0 \\ x \perp e_1, \ldots, x \perp e_{k-1}}} R(x)$$

Proof. By the Real Spectral Theorem we know $e_1, \ldots, e_n$ form a basis of $\mathcal{V} = \mathbb{R}^n$. For any nonzero $x \in \mathcal{V}$ (note $\|x\|^2 = 0$ iff $x = 0$), we can write

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

$$\langle x, Mx \rangle = \left\langle \sum_{i=1}^{n} \langle x, e_i \rangle e_i, \sum_{i=1}^{n} \lambda_i \langle x, e_i \rangle e_i \right\rangle = \sum_{i=1}^{n} \langle x, e_i \rangle \langle e_i, \lambda_i e_i \rangle = \sum_{i=1}^{n} \langle x, e_i \rangle^2 \lambda_i$$

$$\|x\|^2 = \left\langle \sum_{i=1}^{n} \langle x, e_i \rangle e_i, \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\rangle = \sum_{i=1}^{n} \langle x, e_i \rangle^2$$

Thus

$$x = \frac{\langle x, Mx \rangle}{\|x\|^2} = \frac{\sum_{i=1}^{n} \langle x, e_i \rangle^2 \lambda_i}{\sum_{i=1}^{n} \langle x, e_i \rangle^2} \leq \frac{\sum_{i=1}^{n} \langle x, e_i \rangle^2 \lambda_1}{\sum_{i=1}^{n} \langle x, e_i \rangle^2} = \lambda_1$$

We also have

$$\frac{\langle e_1, Me_1 \rangle}{\|e_1\|^2} = \lambda_1$$

so the sup is achieved and thus the max exists. To get the case for general k, we see that since A is symmetric (i.e., $A = A^T$), the spectral norm simplifies to:

$$\|A\|_2 = \max_{1 \leq i \leq n} |\lambda_i|$$

where λ_i are the eigenvalues of A.

Let's go through the proof that for a symmetric matrix A, such as the adjacency matrix of an undirected graph, the spectral norm equals the largest eigenvalue in absolute value.

Let $A = \sum_i \lambda_i v_i v_i^T$, where $\{v_i\}_i$ is an orthonormal eigenbasis of A, as A is symmetric. Then,

$$x^T A x = x^T A \sum_i v_i v_i^T x = x^T \sum_i \lambda_i v_i v_i^T x \leq x^T \max_i \{\lambda_i\} x = \max_i \{\lambda_i\} = \|A\|$$

$\square$

what does spectral norm tell us about the graph?

Analogy: Stretching a Rubber Band

Imagine you have a rubber band shaped like a circle. You apply a transformation using the matrix A, which stretches, squishes, or distorts the rubber band into a new shape.

The **spectral norm** measures:

> What is the maximum amount the rubber band stretches in any direction?

Some directions may stretch a little, while others stretch a lot. The spectral norm captures the *largest possible stretch factor*.

For Adjacency Matrices of Graphs

If A is the adjacency matrix of a graph, the spectral norm measures:

How much influence can the graph structure amplify information?

In a d-regular graph, the spectral norm equals d, meaning the maximum stretching effect corresponds to the degree d, the maximum local connectivity.

Connection to Eigenvalues

The spectral norm for symmetric matrices (like adjacency matrices of undirected graphs) simplifies to:

$$\|A\|_2 = \max_{1 \leq i \leq n} |\lambda_i|$$

This is the *largest eigenvalue in absolute value*, providing a simple yet powerful way to understand how "strong" or "connected" the graph is.

Theorem 4. *Let G be a d-regular graph with adjacency matrix A. Then all eigenvalues of A lie within the interval $[-d, d]$.*

Since A is symmetric, its spectral norm is equal to the largest absolute eigenvalue:

$$\|A\| = \max_{1 \leq i \leq n} |\lambda_i|$$

Thus, proving $\|A\| \leq d$ suffices to show that all eigenvalues of A lie in $[-d, d]$. Step 2: Bounding $x^T A^2 x$ for Any Unit Vector Consider any unit vector x (i.e., $\|x\|_2 = 1$). We examine the quadratic form:

$$x^T A^2 x$$

Expanding in terms of matrix entries,

$$x^T A^2 x = \sum_{i,j} \left(A^2\right)_{ij} x_i x_j$$

For a d-regular graph, each vertex has exactly d neighbors, and the adjacency matrix satisfies:

$$\left(A^2\right)_{ii} = d, \quad \forall i$$

Thus, the sum simplifies to:

$$x^T A^2 x = \sum_{i=1}^{n} d x_i^2 = d \sum_{i=1}^{n} x_i^2 = d$$

Since the largest eigenvalue of A^2 is the square of the largest eigenvalue of A, we obtain:

$$\lambda_{\max}\left(A^2\right) \leq d^2$$

Taking square roots:

$$\lambda_{\max}(A) \leq d$$

Since A is symmetric, its spectrum is symmetric about zero, meaning:

$$\lambda_{\min}(A) \geq -d$$

块对角矩阵的特征值是各块特征值的并集。设每个 A_i 的最大特征值为 μ_i，则整个矩阵 A 的最大特征值为:

$$\lambda_1 = \max_{1 \le i \le k} \mu_i.$$

Thus, all eigenvalues of A lie within $[-d, d]$, completing the proof.

□

A d-regular graph has eigenvalue d corresponding to the eigenvector $\mathbf{1}$ (the all-ones vector), which is the case for any regular graph. This means that the sum of the rows (or columns) of the adjacency matrix is equal to d, so the all-ones vector is a natural eigenvector.

- However, the eigenvalue - d typically does not always appear in all d-regular graphs, except in very specific cases (such as in bipartite graphs).

2. Eigenvalue $-d$ in Bipartite Graphs: - The only case where you are guaranteed to see the eigenvalue $-d$ is in a bipartite d regular graph.

- In a bipartite d-regular graph, the graph is divided into two sets of vertices V_1 and V_2, where each vertex in V_1 is connected to exactly d vertices in V_2, and each vertex in V_2 is connected to exactly d vertices in V_1. There are no edges within V_1 or within V_2. In a non-bipartite d-regular graph, the eigenvalue d does not necessarily exist.

4 Connectivity and Spectral Radius

Let's explore a fascinating property of connected graphs using their *eigenvalues*. Specifically, we're going to focus on the relationship between the *connectivity* of a graph and the *eigenvalues* of its *adjacency matrix*.

Theorem 5. *Let G be an undirected graph with adjacency matrix A. The graph G is connected if and only if the second-largest eigenvalue (in absolute value) is strictly less than the largest eigenvalue.*

Proof. Proof (Outline) Since A is symmetric for undirected graphs, it has a full set of orthonormal eigenvectors. The largest eigenvalue λ_1 (Perron-Frobenius property for connected graphs with nonnegative entries) corresponds to a strictly positive eigenvector. If the graph is connected, all other eigenvectors must be orthogonal to this, meaning they must change sign (some positive, some negative), indicating they describe cuts between subsets of the graph. The spectral gap — the difference between λ_1 and λ_2 — measures how tightly connected the graph is. In particular, if there are disconnected components, there will be eigenvalues equal to λ_1, indicating these components. □

Applications

- **Community Detection:** Large spectral gaps indicate strong global connectivity, while small gaps may indicate clusters or communities.
- **Random Walk Convergence:** Spectral gaps control the mixing time of random walks on graphs, crucial in ranking algorithms like PageRank.
- **Graph Partitioning:** Spectral partitioning algorithms use the second eigenvector (the Fiedler vector) to cut graphs into balanced parts.

- **Application**: Rapid connectivity checks in power grid networks using sparse eigenvalue computations.

- **Extension**: Disconnected graphs have multiple eigenvalues with magnitude λ_1, useful for detecting network fragmentation.

5 Section 5: Cheeger's Inequality and the Adjacency Matrix

Cheeger's inequality connects the spectral properties of a graph, particularly the eigenvalues of its adjacency matrix, to its expansion properties. The key feature of a graph that this inequality highlights is its *edge expansion*, which measures how well the vertices of the graph are connected.

The *Cheeger constant* measures the difficulty of cutting a graph into two roughly equal parts:

$$h(G) = \min_{S \subset V, 0 < |S| \leq \frac{n}{2}} \frac{e(S, \bar{S})}{d|S|}$$

where $e(S, \bar{S})$ counts edges crossing from S to the rest of the graph. It captures how many edges you need to cut to split the graph.

Intuitive Meaning

- If $h(G)$ is large, the graph is very connected — there's no narrow bottleneck. - If $h(G)$ is small, there's an easy way to cut the graph into two parts with few edges between them.

Cheeger's Inequality

Cheeger's Inequality links $h(G)$ to λ_2, the second-largest eigenvalue of the adjacency matrix:

$$\frac{1}{2}(d - \lambda_2) \leq h(G) \leq \sqrt{2d(d - \lambda_2)}$$

This gives a precise numerical relationship between:

- How easy the graph is to cut (captured by $h(G)$).

- How much the graph stretches vectors (captured by λ_2).

Proof. Outline

Regular graphs serve as a useful benchmark or reference model. If we understand how the spectral properties of a regular graph behave, we can more easily compare them to irregular graphs, which might have more complicated structures.

In this part, we assume for simplicity that G is a d-regular graph. We shall work with the normalized adjacency matrix $\mathbf{M} = \frac{1}{d}\mathbf{A}$. The goal of this class is to prove Cheeger's inequality, which establishes an interesting connection between $1 - \lambda_2$ and the (normalized) edge expansion. Here λ_2 is the second largest eigenvalue of M.

key preparations:

$$\sum_{u,v \in V} \mathbf{M}_{u,v} \cdot (x_u - x_v)^2 = 2x^{\mathrm{T}}x - 2x^{\mathrm{T}}\mathbf{M}x$$

(Courant-Fischer Formula for λ_1 and λ_2)

The conductance:

$$\Phi(G) := \min_{\emptyset \subsetneq S \subsetneq V} \frac{|E(S, V \backslash S)|}{d|S| \cdot \frac{|V \backslash S|}{|V|}}$$

Intuitively, the conductance measures how close a graph is to a random d-regular graph. The reason is that for any set S, we have $d|S|$ incident edges and the proportion of edges going to $|V \backslash S|$ is $|V \backslash S|/|V|$.

$$h(G) \leqslant \Phi(G) \leqslant 2h(G)$$

□

Applications of Cheeger's Inequality

Cheeger's inequality has several important applications, particularly in spectral graph theory and network analysis:

- Graph Partitioning: Cheeger's inequality is widely used in spectral clustering algorithms. It helps to identify the "best" way to partition a graph into clusters, ensuring that the clusters are well-separated.
- Random Walks: In the context of random walks on graphs, the second-largest eigenvalue λ_2 determines how quickly a random walk on the graph will converge to a steady state. A large λ_2 implies faster convergence.
- Network Robustness: In network theory, Cheeger's inequality helps to assess the robustness of a network. A graph with a small λ_2 may indicate a vulnerability to disconnection if certain edges or nodes are removed.

By relating the connectivity of a graph to the spectral properties of its adjacency matrix, Cheeger's inequality gives us an invaluable tool for understanding the structural properties of networks, whether in social networks, biological networks, or communication networks.

6 Graph Energy and Chemical Applications

The **graph energy** is defined as:

$$\mathcal{E}(G) = \sum_{i=1}^{n} |\lambda_i|.$$

- **Chemical Connection**: In chemical graph theory, $\mathcal{E}(G)$ approximates the π-electron energy of hydrocarbon molecules.
  Example: Benzene (C_6H_6) corresponds to a hexagonal graph with energy proportional to stability.