1 Normalized Laplacian Matrix

Some of the explanations, derivations, and examples were written directly on the board during the lecture and are not fully included in this document.

In a previous lecture, we saw that in graph theory, especially when dealing with random walks or Markov chains, the transition matrix P is often defined as $D^{-1}A$, where A is the adjacency matrix and D is the diagonal degree matrix. Each entry P_{ij} represents the probability of moving from node i to node j, which is just the weight of the edge (i, j) divided by the degree of node i. That makes sense because the rows of P sum to 1, as required for a transition matrix. However, P is not symmetric, then perhaps people symmetrize it using the degree matrices.

$$P = D^{-1}A = D^{-1/2}\left(D^{-1/2}AD^{-1/2}\right)D^{1/2},$$

so that the eigenvalues of P are the same as the eigenvalues of M, a symmetric matrix. We can use properties of symmetric matrices to help us find and evaluate the eigenvalues of P, which determine the rate of convergence of a random walk. Besides, M is symmetric, ensuring real eigenvalues and orthogonal eigenvectors. Eigenvectors of P are scaled versions of S's eigenvectors by $D^{+-1/2}$.

Our main tool will be the Laplacian, which is defined by

$$\mathcal{L} = I - M = I - D^{-1/2}AD^{-1/2} = D^{-1/2}(D - A)D^{-1/2}.$$

Note that $\left(D^{-1/2}AD^{-1/2}\right)$ is called the normalized adjacency matrix.

$L = D - A$ is referred to as the combinatorial Laplacian, while the Laplacian we have introduced here is referred to as the normalized Laplacian.

From the definition of the Laplacian, we can compute it entrywise.

$$\mathcal{L}(u, v) = \begin{cases} 1 & \text{if } u = v\text{and u is nor isolated;} \\ -\frac{1}{\sqrt{d_u d_v}} & \text{if } u \sim v; \\ 0 & \text{otherwise.} \end{cases}$$

So there are three types of entries. Note when $d_u = 0$, by convention, $\mathcal{L}(u, u) = 0$. Note that we must assume there are no isolated vertices. (Those are uninteresting from a spectral perspective anyway.)

The quadratic form of the normalized Laplacian is derived as follows, for any vector x, the quadratic form is:

$$x^T\mathcal{L}x = x^T(I - D^{-1/2}AD^{-1/2})x.$$

Expanding this:

$$x^T\mathcal{L}x = x^TIx - x^TD^{-1/2}AD^{-1/2}x.$$

Since $x^TIx = \sum_u x_u^2$, and

$$x^TD^{-1/2}AD^{-1/2}x = \sum_{u,v} A_{uv}\frac{x_u x_v}{\sqrt{d_u d_v}},$$

we obtain:

$$x^T \mathcal{L} x = \sum_u x_u^2 - \sum_{u,v} A_{uv} \frac{x_u x_v}{\sqrt{d_u d_v}}.$$

Using the identity:

$$x_u^2 + x_v^2 - 2x_u x_v = (x_u - x_v)^2,$$

we rewrite the expression as:

$$x^T \mathcal{L} x = \frac{1}{2} \sum_{u \sim v} A_{uv} \left(\frac{x_u^2}{\sqrt{d_u d_v}} + \frac{x_v^2}{\sqrt{d_u d_v}} - 2\frac{x_u x_v}{\sqrt{d_u d_v}} \right).$$

Since $A_{uv} = 1$ for adjacent vertices, this simplifies to:

$$x^T \mathcal{L} x = \sum_{u \sim v} (\frac{x_u}{\sqrt{d_u}} - \frac{x_v}{\sqrt{d_v}})^2$$

Since the Laplacian is symmetric, all the eigenvalues are real; further, it is easy to show that the Laplacian is positive semi-definite (i.e., all the eigenvalues are non-negative). We let

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_n$$

denote the eigenvalues of the Laplacian.

Okay, now the question comes, does $\mathcal{L}$ behave more like A or L? The answer will be "kind of".

1.1 Relation with Adjacency matrix

First, we look at ways in which the normalized Laplacian is related to the adjacency matrix. Recall that for the largest eigenvalue of A and the maximum degree of a vertex in a graph, davg . "Normalizing" the adjacency matrix (degree-based scaling) makes its largest eigenvalue

Normalizing the adjacency matrix ensures its largest eigenvalue becomes 1 due to degree-based scaling and properties of symmetric matrices. Here's the breakdown: 1. Normalization Process The symmetric normalized adjacency matrix is defined as:

$$S = D^{-1/2} A D^{-1/2},$$

where D is the degree matrix (diagonal matrix of node degrees) and A is the adjacency matrix. This normalization scales A by the square roots of the degrees of its nodes.

2. Key Eigenvalue Properties Eigenvalue 1 via the Degree Vector: Consider the vector $\mathbf{v} = D^{1/2}\mathbf{1}$, where $\mathbf{1}$ is a vector of all ones. Applying S to $\mathbf{v}$:

$$S\mathbf{v} = D^{-1/2} A D^{-1/2} \cdot D^{1/2}\mathbf{1} = D^{-1/2} A \mathbf{1}.$$

Since $A\mathbf{1} = D\mathbf{1}$ (the adjacency matrix acting on $\mathbf{1}$ returns the degree vector),

$$S\mathbf{v} = D^{-1/2} D \mathbf{1} = D^{1/2}\mathbf{1} = \mathbf{v}.$$

Thus, $\mathbf{v}$ is an eigenvector of S with eigenvalue 1.

- Maximum Eigenvalue is 1: The spectral radius (largest eigenvalue) of S is bounded by its operator norm. For symmetric matrices, the spectral radius equals the maximum

absolute eigenvalue. Since S is a contraction (entries are scaled by degrees), its eigenvalues lie in $[-1, 1]$. However, because S is non-negative and irreducible (for connected graphs), the Perron-Frobenius theorem guarantees that 1 is the largest eigenvalue.

3. Intuition Behind Normalization - Scaling by Degrees: Normalizing A with $D^{-1/2}$ ensures no row/column of S has a norm greater than 1. This prevents high-degree nodes from dominating the spectrum. - Connection to Stochastic Matrices: The random walk transition matrix $P = D^{-1}A$ has a largest eigenvalue of 1 (as it's stochastic). Since S is similar to P, they share eigenvalues, including the dominant eigenvalue 1.

4. Example: Regular Graphs In a d-regular graph (all nodes have degree d): - $D = dI$, so $S = \frac{1}{d}A$. - The largest eigenvalue of A is d, so scaling by $\frac{1}{d}$ gives $\lambda_{\max}(S) = 1$.

For non-regular graphs, normalization ensures eigenvalues are scaled proportionally to degrees, but 1 remains the largest eigenvalue due to the eigenvector $\mathbf{v} = D^{1/2}\mathbf{1}$.

5. Why Does This Matter? Stability: Bounding eigenvalues prevents numerical issues in algorithms (e.g., graph neural networks).

Interpretability: Eigenvalues in $[0, 1]$ relate to diffusion processes (e.g., random walks "settle" to stationary distributions).

Spectral Analysis: Symmetry ensures real eigenvalues, simplifying clustering and embedding techniques.

Summary: Normalizing A to $S = D^{-1/2}AD^{-1/2}$ scales the matrix so that its largest eigenvalue is 1, anchored by the eigenvector encoding node degrees. This reflects the balanced influence of nodes in the graph, regardless of degree heterogeneity.

Let us see in another way: Have that when $A\vec{x} = \lambda\vec{x}$, then for every vertex v we have

$$\sum_{u,\, uv \in E(G)} x_u = \lambda x_v.$$

If we make A weighted, then

Now, suppose that $\mathcal{L}\vec{x} = \lambda\vec{x}$.

The eigenvalue equation for the normalized Laplacian is:

$$\mathcal{L}x = \lambda x.$$

Substituting the definition of $\mathcal{L}$:

$$(I - D^{-1/2}AD^{-1/2})x = \lambda x.$$

Rearranging, we get:

$$(1 - \lambda)x = D^{-1/2}AD^{-1/2}x.$$

Transformation to Harmonic Eigenvector Multiplying both sides by $D^{1/2}$, we obtain:

$$(1 - \lambda)D^{1/2}x = AD^{-1/2}x.$$

We let $\vec{y} = D^{-\frac{1}{2}}\vec{x}$. We call this vector a *harmonic eigenvector*, and note that it is not an eigenvector of $\mathcal{L}$ although it is related to one.

we get:

$$(1-\lambda)Dy = Ay.$$

This equation shows that y satisfies a weighted sum condition at each vertex, making it a **harmonic-like function**.

for every vertex v, we have

$$\sum_{u,\, uv \in E(G)} y_u = (1-\lambda)d_v y_v.$$

Weighted version:

$$\sum_{u \sim v} w(u,v) y_u = (1-\lambda)d_v y_v \tag{1}$$

shows that y satisfies a weighted averaging property, meaning it resembles **harmonic functions** in classical analysis.

The spectrum of $\mathcal{L}$ lies between 0 and 2. The multiplicity of the eigenvalue 0 corresponds to the number of connected components in the graph. *The multiplicity of the eigenvalue 0 of A is equal to the multiplicity of the eigenvalue 1 of $\mathcal{L}$.*

The spectrum of $\mathcal{L}$ lies between 0 and 2. The multiplicity of the eigenvalue 0 corresponds to the number of connected components in the graph.

Show that

$$0 \leq \frac{\vec{x}^T L \vec{x}}{\vec{x}^T \vec{x}} \leq 2.$$

What do the multiplicities of the eigenvalues 0 and 2 tell us about G?

The bipartite graph has the largest value 2 and the multiplicities of 2 as eigenvalue for the normalized laplacian matrix is equal to the number of bipartite components.

recall that for a bipartite graph. if λ is an eigenvalue of its adjacency matrix A if and only if λ is, with multiplicity; that is, if the multiplicity of λ is k, then so is the multiplicity of $(-\lambda, (v1, -v2)))$.

Further. Suppose G is connected. Then, $1 = n$ if and only if G is bipartite. for complete bipartite graph.,

$$-\sqrt{mn}, 0, ..., 0, \sqrt{mn}.$$

Eigenvalues in the range $[0, 1]$ of $\mathcal{L}$ correspond to the positive eigenvalues of the adjacency matrix A; the other to negative eigenvalues of A.

Following the normalized idea, $I = D^{-1/2}D^{1/2}$. The normalized Laplacian essentially captures how much a node differs from its normalized adjacent neighbors.

The following table summarizes some aspects of graphs that can be deduced from the spectra of various associated matrices.

	nr. components	*bipartite?*	*nr. bip components*	*nr. edges*
A	*no*	*yes*	*no*	*yes*
L	*yes*	*no*	*no*	*yes*
Q	*no*	*no*	*yes*	*yes*
$\mathcal{L}$	*yes*	*yes*	*yes*	*no*