# Eigenvalues of Mixing, Incidence, Adjacency, and Laplacian Matrices of an Undirected Graph

Consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices with cardinality $n$ and $\mathcal{E}$ is the set of edges with cardinality $m$. The graph does not contain self-loops. We are interested in the mixing, incidence, adjacency, and Laplacian matrices of $\mathcal{G}$, and will investigate the properties of their eigenvalues.

**Definition 1 (Stochastic Matrix)**. Let $P \in \mathbb{R}^{n\times n}$ be a nonnegative matrix whose entries $p_{ij} \geq 0$. $P$ is said to be row-stochastic (right-stochastic) the sum of every row is 1, namely, $\sum_{j=1}^{n} p_{ij} = 1$ for all $i \in [n]$. $P$ is said to be column-stochastic (left-stochastic) the sum of every column is 1, namely, $\sum_{i=1}^{n} p_{ij} = 1$ for all $j \in [n]$. $P$ is said to be doubly stochastic if it is both row-stochastic and column-stochastic.

**Definition 2 (Primitive Matrix) [3]**. Let $P \in \mathbb{R}^{n\times n}$ be a nonnegative matrix whose entries $p_{ij} \geq 0$. $P$ is said to be primitive if there exists an integer $k$ such that $P^k$ is a positive matrix whose entries $(P^k)_{ij} > 0$.

**Definition 3 (Irreducible Matrix) [3]**. Let $P \in \mathbb{R}^{n\times n}$ be a nonnegative matrix whose entries $p_{ij} \geq 0$. $P$ is said to be irreducible if for any $i, j \in [n]$ there exists an integer $k$ such that $(P^k)_{ij} > 0$.

A direct implication is that a primitive matrix is irreducible matrix, but not vise versa.

**Example 1 (Irreducible But Not Primitive) [3]**. A matrix that is irreducible, but not primitive. Consider a matrix

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Observe that $P^k = P$ when $k$ is odd, while $P^k = I$ when $k$ is even.

**Definition 4 (Spectral Radius)**. The spectral radius of a matrix $P \in \mathbb{R}^{n\times n}$ is defined as $\rho(P) := \max_i |\lambda_i(P)|$ where $\lambda_i(P)$ denotes the $i$-th eigenvalue of $P$. Note that $\lambda_i(P)$ may be not a real value.

**Theorem 1 (Perron Theorem) [3]**. For a primitive matrix $P \in \mathbb{R}^{n\times n}$ with spectral radius $\rho(P)$, we have the following conclusions. (i) $\rho(A) > 0$ and $\rho(A)$ is an eigenvalue of $P$ with multiplicity 1. (ii) The left and right eigenvectors of $P$ corresponding to the eigenvalue $\rho(A)$ are both positive. (iii) For any other eigenvalue $\lambda$ of $P$, $|\lambda| < \rho(P)$.

**Theorem 2 (Perron-Frobenius Theorem) [3]**. For a irreducible matrix $P \in \mathbb{R}^{n\times n}$ with spectral radius $\rho(P)$, we have the following conclusions. (i) $\rho(A) > 0$ and $\rho(A)$ is an eigenvalue of $P$ with multiplicity 1. (ii) The left and right eigenvectors of $P$ corresponding to the eigenvalue $\rho(A)$ are both positive.

**Example 2 (Perron Theorem Versus Perron-Frobenius Theorem)**. Consider a matrix that is irreducible, but not

primitive, given by

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Its eigenvalues are $-1$ and $1$, meaning that (iii) in Theorem 1 does not hold any more.

**Theorem 3 (Gershgorin Circle Theorem) [3]**. Any eigenvalue $\lambda$ of a matrix $P \in \mathbb{R}^{n\times n}$ must lie in at least one of the discs

$$|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|, \ i = 1, 2, \cdots, n.$$

**Property 1 (Largest and Smallest Eigenvalues of Row-Stochastic Matrix)**. If $P \in \mathbb{R}^{n\times n}$ is a row-stochastic matrix, any eigenvalue $\lambda$ must lie in at least one of the discs $[2a_{ii} - 1, 1]$. It is obvious that $P$ has at least one eigenvalue at 1. By Gershgorin circle theorem, the largest eigenvalue of $P$ is also 1. If further $a_{ii} > 0$ for all $i \in [n]$, then the smallest eigenvalue is larger than $-1$.

**Property 2 (Markov Chain and State Transition Probability) [3]**. Suppose that a row-stochastic matrix $P \in \mathbb{R}^{n\times n}$ describes a Markov chain. It is obvious that $P$ has at least one eigenvalue at 1. By Gershgorin circle theorem, the largest eigenvalue of $P$ is also 1. There might be multiple left and right eigenvectors corresponding to eigenvalue 1. By $P\frac{1_n}{n} = \frac{1_n}{n}$ we know $\frac{1_n}{n}$ is a corresponding right eigenvector.

Consider a $n$-dimensional row probability vector $a^T$ whose elements $a_i$ are nonnegative and summed to 1, namely, $a_i \geq 0$ for all $i \in [n]$ and $\sum_{i=1}^n a_i = 1$. Let $a_i$ define the probability of state $i$ at time 0. Then, the probability of state $j$ at time $k$ is $\sum_{i=1}^n a_i (P^k)_{ij}$. A $n$-dimensional row probability vector $\pi^T$ is called as stationary if it satisfies $\pi^T = \pi^T P$, namely, $\pi^T$ is a left eigenvector corresponding to eigenvalue 1. Note that such a row probability vector may not exist. If it does exist, apparently, $\pi^T = \pi^T P^k$ for all $k = 0, 1, \cdots$.

If the row-stochastic matrix $P$ is irreducible, then there exist at least one such stationary row probability vector $\pi^t$. Further, if $P$ is primitive, then the stationary row probability vector $\pi^T$ is unique and its $j$-th element is given by $\pi_j = \lim_{k\to\infty}(P^k)_{ij}$. That is, every row of $\lim_{k\to\infty} P^k$ is the row probability vector $\pi^T$. Recall that the limiting state transition probability from state $i$ to state $j$ given by

$$\lim_{k\to\infty} P(s^k = j | s^0 = i) = \lim_{k\to\infty} (P^k)_{ij}.$$

Then in this case, we know that the limiting state transition probability is irrelevant with the initial state $i$.

**Definition 5 (Mixing Matrix)**. A nonnegative mixing matrix $W \in \mathbb{R}^{n\times n}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ has the following properties. (i) Decentralized computation. $W \geq 0$. If $(i, j) \notin \mathcal{E}$ and $i \neq j$, then $w_{ij} = 0$. (ii) Stochastic matrix. $W$ is at least row-stochastic such that $W1_n = 1_n$, and often doubly stochastic such that $W1_n = 1_n$ and $W = W^T$. (iii) Proper eigenvalues. The largest eigenvalue is 1 and with multiplicity 1. Often, it is also required that the smallest eigenvalue is larger than $-1$.

Since $W$ is row-stochastic, by Gershgorin Circle Theorem the spectral radius $\rho(W) = 1$. If $W$ is irreducible, by Perron-Frobenius Theorem, 1 is an eigenvalue of $W$ with multiplicity 1. If $W$ is further primitive, by Perron

Theorem, 1 is an eigenvalue of $W$ with multiplicity 1 and for any other eigenvalue $\lambda$ of $W$, $|\lambda| < 1$.

**Example 3 (Mixing Matrices with Eigenvalues in $[-1, 1]$) [2]**. Denote $d_i$ as the degree of node $i$ and $d_{\max} := \max_{i \in \mathbf{V}} d_i$ as the maximum degree. The maximum-degree rule generates the mixing matrix $W$ by

$$w_{ij} = \begin{cases} \frac{1}{d_{\max}} & \text{if } (i,j) \in \mathcal{E}, \\ 1 - \frac{d_i}{d_{\max}} & \text{if } i = j, \\ 0 & \text{if } (i,j) \notin \mathcal{E} \text{ and } i \neq j. \end{cases} \tag{1}$$

For bipartite graphs, $W$ generated by the maximum-degree rule has at least one eigenvalue at $-1$. An example is a connected two-node network, which has

$$W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The two eigenvalues are $-1$ and 1.

The Metropolis-Hastings rule generates the mixing matrix $W$ by

$$w_{ij} = \begin{cases} \min\{\frac{1}{d_i}, \frac{1}{d_j}\} & \text{if } (i,j) \in \mathcal{E}, \\ \sum_{(i,k) \in \mathcal{E}} \max\{0, \frac{1}{d_i} - \frac{1}{d_k}\} & \text{if } i = j, \\ 0 & \text{if } (i,j) \notin \mathcal{E} \text{ and } i \neq j. \end{cases} \tag{2}$$

According to the Metropolis-Hastings rule, for a connected two-node network, the generated $W$ is also

$$W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The two eigenvalues are $-1$ and 1.

**Example 4 (Mixing Matrices with Eigenvalues in $(-1, 1]$) [5]**. The Metropolis constant edge weight rule generates $W$ as

$$w_{ij} = \begin{cases} \frac{1}{\max\{d_i, d_j\} + \epsilon} & \text{if } (i,j) \in \mathcal{E}, \\ 1 - \sum_{k \in \mathcal{V}} w_{ik} & \text{if } i = j, \\ 0 & \text{if } (i,j) \notin \mathcal{E} \text{ and } i \neq j. \end{cases} \tag{3}$$

Here $\epsilon > 0$ is a small positive constant.

The Laplacian-based constant edge weight rule generates $W$ as

$$W = I - \frac{L}{\tau}. \tag{4}$$

Here $L$ is the oriented Laplacian matrix and $\tau > \frac{1}{2}\lambda_{\max}(L)$ is a scaling factor. When $\lambda_{\max}(L)$ is not available, $\tau = d_{\max} + \epsilon$ for some small $\epsilon$, say $\epsilon = 1$, can be used.

Given a bidirectionally connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, it is shown in [4] that if the constructed mixing matrix $W$ is row-stochastic, $w_{ij} > 0$ if $(i,j) \in \mathcal{E}$, and there exists at least one $a_{ii} > 0$, then $W$ is primitive. Since $W$ is row-stochastic, by Gershgorin circle Theorem the spectral radius $\rho(W) = 1$. Since $W$ is primitive, by Perron

Theorem, 1 is an eigenvalue of $W$ with multiplicity 1 and for any other eigenvalue $\lambda$ of $W$, $|\lambda| < 1$.

For more options of mixing matrices, see Table 14.1 in [4].

**Definition 6 (Incidence Matrix)**. An unoriented incidence matrix $\tilde{B} \in \mathbb{R}^{n\times m}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed as follows. For an undirected edge $e = (i, j)$, $\tilde{b}_{ie} = 1$ and $\tilde{b}_{je} = 1$. An (oriented) incidence matrix is often associated with a directed graph, but here we adapt it to describe an undirected graph. An (oriented) incidence matrix $B \in \mathbb{R}^{n\times m}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed as follows. For an undirected edge $e = (i, j)$, if $i < j$, then $b_{ie} = 1$ and $b_{je} = -1$; otherwise if $j < i$, then $b_{ie} = -1$ and $b_{je} = 1$.

**Definition 7 (Adjacency Matrix)**. An adjacency matrix $A \in \mathbb{R}^{n\times n}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed as follows. If $(i, j) \in \mathcal{E}$, then $a_{ij} = 1$ and $a_{ji} = 1$; otherwise if $(i, j) \notin \mathcal{E}$, then $a_{ij} = 0$ and $a_{ji} = 0$.

**Definition 8 (Degree Matrix)**. An degree matrix $D \in \mathbb{R}^{n\times n}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a diagonal matrix whose $i$-th diagonal entry $d_i$ is the degree of node $i$.

**Definition 9 (Laplacian Matrix)**. A signless Laplacian matrix $L \in \mathbb{R}^{n\times n}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is $\tilde{L} = D+A$. A (signed) Laplacian matrix $L \in \mathbb{R}^{n\times n}$ associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is $L = D - A$. For (signed) Laplacian matrix $L$ and (oriented) incidence matrix $B$, we have $L = BB^T$.

The symmetric normalized Laplacian matrix is defined as $L^{sym} := D^{-1/2}LD^{-1/2} = I - D^{-1/2}AD^{-1/2}$. The elements of $L^{sym}$ are given by

$$L_{ij}^{sym} = \begin{cases} -\frac{1}{\sqrt{d_i d_j}} & \text{if } (i,j) \in \mathcal{E}, \\ 1 & \text{if } i = j, \\ 0 & \text{if } (i,j) \notin \mathcal{E} \text{ and } i \neq j. \end{cases} \tag{5}$$

The random-walk normalized Laplacian matrix is defined as $L^{rw} := D^{-1}L = I - D^{-1}A$. The elements of $L^{rw}$ are given by

$$L_{ij}^{rw} = \begin{cases} -\frac{1}{d_i} & \text{if } (i,j) \in \mathcal{E}, \\ 1 & \text{if } i = j, \\ 0 & \text{if } (i,j) \notin \mathcal{E} \text{ and } i \neq j. \end{cases} \tag{6}$$

The name The name of random-walk normalized Laplacian comes from the fact that $L^{rw} = I-P$, where $P = D^{-1}A$ is simply the transition matrix of a random walker on the graph.

**Property 3 ((Signed) Laplacian Matrix) [6]**. The (signed) Laplacian matrix $L$ is symmetric as $\mathcal{G}$ is bidirectionally connected. Therefore, all the eigenvalues of $L$ are real values. We sort the eigenvalues in an ascending order, as $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

(i) $L$ is positive semidefinite such that $\lambda_1 \geq 0$. This comes from that $L$ is symmetric and diagonally dominant. An alternative viewpoint is that $L = BB^T$. Another alternative viewpoint is by Gershgorin circle Theorem.

(ii) $\lambda_1 = 0$ because every row sum and column sum of $L$ is 0 such that $L\frac{1_n}{n} = 0 \cdot \frac{1_n}{n}$.

(iii) The number of connected components in $L$ is the dimension of null$(L)$ and the multiplicity of eigenvalue 0. Because we assume that $\mathcal{G}$ is connected, null$(L) =$ span$(1_n)$.

(iv) The smallest non-zero eigenvalue of $L$ is called the spectral gap. Because we assume that $\mathcal{G}$ is connected, the spectral gap is larger than 0.

(v) The second smallest eigenvalue of $L$ (could be 0) is the algebraic connectivity (or Fiedler value) of $\mathcal{G}$ and approximates the sparsest cut of $\mathcal{G}$.

(vi) By Gershgorin circle Theorem, all eigenvalues of (signed) Laplacian matrix $L$ is within $[0, 2d_{\max}]$ where $d_{\max} := \max_i d_i$ is the maximum degree. All eigenvalues of symmetric normalized Laplacian matrix $L^{sym}$ is within $[0, 2]$. All eigenvalues of random-walk normalized Laplacian matrix $L^{rw}$ is within $[0, 2]$.

**Property 4 (Symmetric Normalized Laplacian) [1]**. We sort the eigenvalues of symmetric normalized matrix $L^{sym}$ in an ascending order, as $\lambda_1^{sym} \leq \lambda_2^{sym} \leq \cdots \leq \lambda_n^{sym}$.

(i) $\sum_{i=1}^n \lambda_i^{sym} \leq n$ and the equality holds if and only if $\mathcal{G}$ has no isolated nodes. This conclusion comes from observing the trace of $L^{sym}$.

(ii) For $n \geq 2$, $\lambda_2^{sym} \leq \frac{n}{n-1}$ and the equality holds if and only if $\mathcal{G}$ is a complete graph. Also, for $\mathcal{G}$ without isolated nodes, $\lambda_n^{sym} \geq \frac{n}{n-1}$. For $\mathcal{G}$ which is not complete, $\lambda_2^{sym} \leq 1$.

(iii) When $\mathcal{G}$ is bidirectionally connected, $\lambda_2^{sym} > 0$. If $\lambda_i = 0$ and $\lambda_{i+1} \neq 0$, then $\mathcal{G}$ has exactly $i$ connected components.

## REFERENCES


[1] F. R. K. Chung, *Spectral Graph Theory*, American Mathematical Society, 1997

[2] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing Markov chain on a graph," SIAM Review, vol. 46, no. 4, pp. 667–689, 2004

[3] S. U. Pillai, T. Suel, and S. Cha, "The Perron-Frobenius theorem: Some of its applications," IEEE Signal Processing Magazine, vol. 22, no. 2, pp. 62–75, 2005

[4] A. H. Sayed, "Adaption, learning, and optimization over networks," Foundations and Trends in Machine Learning, vol. 7, no. 4–5, pp. 311–801, 2014

[5] W. Shi, Q. Ling, G. Wu, and W. Yin, "EXTRA: An exact first-order algorithm for decentralized consensus optimization," SIAM Journal on Optimization, vol. 25, no. 2, pp. 944–966, 2015

[6] https://en.wikipedia.org/wiki/Laplacian_matrix