

Spectral Clustering Algorithms

Learner

L^AT_EX by Xin Wang

Notes

Contents

1	Graph theory	4
1.1	Basic terminologies	4
1.2	Graphs	5
1.3	Edges	6
1.4	Matrices	7
1.4.1	Symmetric matrix	7
1.4.2	Standard basis	10
1.4.3	Diagonal matrix	11
1.4.4	Degree matrix	14
1.4.5	Adjacency matrix	14
1.4.6	Weighted Adjacency matrix	15
1.4.7	Weighted vertex-edge matrix	16
1.4.8	Normalized weighted vertex-edge matrix	17
1.4.9	Laplacian matrix	19
1.4.10	Normalized Laplacian matrix	21
1.4.11	Eigendecomposition	24
1.4.12	Matrix Exponential	27
1.4.13	QR decomposition	29
2	Classical Spectral Clustering Algorithm	33
2.1	Classical Spectral Clustering Algorithm	33

List of Figures

List of Tables

Chapter 1

Graph theory

1.1 Basic terminologies

Definition 1.1.1: Network or Graph

A network (or graph) consists of a set of nodes (or vertices, actors) and a set of edges (or links, ties) that connect those nodes[1].

Definition 1.1.2: Neighbor

Node j is called a neighbor of node i if (and only if) node i is connected to node j .

Definition 1.1.3: Degree

The number of edges connected to a node. Node i 's degree is often written as $\deg(i)$.

Definition 1.1.4: Walk

A list of edges that are sequentially connected to form a continuous route on a network.

Definition 1.1.5: Trail

A walk that doesn't go through any edge more than once.

Definition 1.1.6: Path

A walk that doesn't go through any node (and therefore any edge, too) more than once.

Definition 1.1.7: Cycle

A walk that starts and ends at the same node without going through any node more than once on its way.

Definition 1.1.8: Subgraph

Part of the graph.

Definition 1.1.9: Connected graph

A graph in which a path exists between any pair of nodes.

Definition 1.1.10: Connected component

A subgraph of a graph that is connected within itself but not connected to the rest of the graph.

1.2 Graphs

Definition 1.2.1: Complete graph

A graph in which any pair of nodes are connected.

Definition 1.2.2: Regular graph

A graph in which all nodes have the same degree. Every complete graph is regular.

Definition 1.2.3: Bipartite (n-partite) graph

A graph whose nodes can be divided into two (or n) groups so that no edge connects nodes within each group.

Definition 1.2.4: Tree graph

A graph in which there is no cycle. A graph made of multiple trees is called a forest graph. Every tree or forest graph is bipartite.

Definition 1.2.5: Planar graph

A graph that can be graphically drawn in a two-dimensional plane with no edge crossings. Every tree or forest graph is planar.

1.3 Edges

Definition 1.3.1: Undirected edge

A symmetric connection between nodes. If node i is connected to node j by an undirected edge, then node j also recognizes node i as its neighbor. A graph made of undirected edges is called an undirected graph. The Adjacency matrix of an undirected graph is always symmetric.

Definition 1.3.2: Directed edge

An asymmetric connection from one node to another. Even if node i is connected to node j by a directed edge, the connection isn't necessarily reciprocated from node j to node i . A graph made of directed edges is called a directed graph. The Adjacency matrix of a directed graph is generally asymmetric.

Definition 1.3.3: Unweighted edge

An edge without any weight value associated to it. There are only two possibilities between a pair of nodes in a network with unweighted edges; whether there is an edge between them or not. The Adjacency matrix of such a network is made of only 0's and 1's.

Definition 1.3.4: Weighted edge

An edge with a weight value associated to it. A weight is usually given by a non-negative real number, which may represent a connection strength or distance between nodes, depending on the nature of the system being modeled. The definition of the Adjacency matrix can be extended to contain those edge weight values for networks with weighted edges. The sum of the weights of edges connected to a node is often called the node strength, which corresponds to a node degree for unweighted graphs.

Definition 1.3.5: Multiple edges

Edges that share the same origin and destination. Such multiple edges connect two nodes more than once.

Definition 1.3.6: Self-loop

An edge that originates and ends at the same node.

Definition 1.3.7: Simple graph

A graph that doesn't contain directed, weighted, or multiple edges, or self-loops. Traditional graph theory mostly focuses on simple graphs.

Definition 1.3.8: Multigraph

A graph that may contain multiple edges. Many mathematicians also allow multigraphs to contain self-loops. Multigraphs can be undirected or directed.

Definition 1.3.9: Hyperedge

A generalized concept of an edge that can connect any number of nodes at once, not just two. A graph made of hyperedges is called a hypergraph (not covered in this textbook).

1.4 Matrices

1.4.1 Symmetric matrix

Definition 1.4.1: Symmetric matrix

A real matrix is called symmetric if $\mathbf{A}^T = \mathbf{A}$.
Symmetric matrices are also called selfadjoint.
For complex matrices we would ask $\mathbf{A}^* = \overline{\mathbf{A}}^T = \mathbf{A}$.

Corollary 1.4.2

A symmetric matrix has real eigenvalues.

Proof. Extend the dot product to complex vectors by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \mathbf{v}_i \overline{\mathbf{w}_i}$$

, where $\overline{\mathbf{w}_i}$ is the complex conjugate of \mathbf{w}_i .

For real vectors it is the usual dot product $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \mathbf{v}_i \mathbf{w}_i = \mathbf{v} \cdot \mathbf{w}$.

For a real matrix \mathbf{A} , we have

$$\langle \mathbf{A}\mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n (\mathbf{A}\mathbf{v})_i \overline{\mathbf{w}_i} \tag{1.1}$$

$$= \sum_{i=1}^n \left(\sum_{j=1}^n \mathbf{A}_{ij} v_j \right) \overline{\mathbf{w}_i} \tag{1.2}$$

$$= \sum_{i=1}^n \sum_{j=1}^n \mathbf{A}_{ij} v_j \overline{\mathbf{w}_i} \tag{1.3}$$

Since \mathbf{A} is a real matrix, then \mathbf{A}_{ij} is real, therefore $\overline{\mathbf{A}_{ij}} = \mathbf{A}_{ij}$.

We can write the j -th component of the complex conjugate of the product of matrix \mathbf{A} and vector \mathbf{w} as:

$$\overline{(\mathbf{A}\mathbf{w})_j} = \sum_{i=1}^n \overline{\mathbf{A}_{ji}\mathbf{w}_i}$$

When we compute the j -th component of the complex conjugate of the product of matrix \mathbf{A}^T and vector \mathbf{w} , we get:

$$\overline{(\mathbf{A}^T\mathbf{w})_j} = \sum_{i=1}^n \overline{(\mathbf{A}^T)_{ji}\mathbf{w}_i}$$

Now, consider the transpose of matrix \mathbf{A} , denoted as \mathbf{A}^T . The element $(\mathbf{A}^T)_{ji}$ of the transpose matrix is equal to the element \mathbf{A}_{ij} of the original matrix. Therefore, we have:

$$\mathbf{A}_{ij} = (\mathbf{A}^T)_{ji}$$

then we obtain:

$$\langle \mathbf{A}\mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \sum_{j=1}^n \mathbf{A}_{ij} \mathbf{v}_j \overline{\mathbf{w}_i} \quad (1.4)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \overline{\mathbf{A}_{ij}} \mathbf{v}_j \overline{\mathbf{w}_i} \quad (1.5)$$

$$= \sum_{j=1}^n \mathbf{v}_j \left(\sum_{i=1}^n \overline{\mathbf{A}_{ij}} \overline{\mathbf{w}_i} \right) \quad (1.6)$$

$$= \sum_{j=1}^n \mathbf{v}_j \left(\sum_{i=1}^n \overline{\mathbf{A}_{ij}\mathbf{w}_i} \right) \quad (1.7)$$

$$= \sum_{j=1}^n \mathbf{v}_j \left(\sum_{i=1}^n \overline{(\mathbf{A}^T)_{ji}\mathbf{w}_i} \right) \quad (1.8)$$

$$= \sum_{j=1}^n \mathbf{v}_j \overline{(\mathbf{A}^T\mathbf{w})_j} \quad (1.9)$$

$$= \langle \mathbf{v}, \mathbf{A}^T\mathbf{w} \rangle \quad (1.10)$$

And assume that the symmetric matrix \mathbf{A} has a complex eigenvalue λ , then

$$\langle \lambda\mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \lambda \mathbf{v}_i \overline{\mathbf{w}_i} \quad (1.11)$$

$$= \lambda \sum_{i=1}^n \mathbf{v}_i \overline{\mathbf{w}_i} \quad (1.12)$$

$$= \lambda \langle \mathbf{v}, \mathbf{w} \rangle \quad (1.13)$$

as well as

$$\langle \mathbf{v}, \lambda \mathbf{w} \rangle = \sum_{i=1}^n \mathbf{v}_i \overline{\lambda \mathbf{w}_i} \quad (1.14)$$

$$= \sum_{i=1}^n \mathbf{v}_i \overline{\lambda} \overline{\mathbf{w}_i} \quad (1.15)$$

$$= \overline{\lambda} \sum_{i=1}^n \mathbf{v}_i \overline{\mathbf{w}_i} \quad (1.16)$$

$$= \overline{\lambda} \langle \mathbf{v}, \mathbf{w} \rangle \quad (1.17)$$

Now, we have:

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \langle \lambda \mathbf{v}, \mathbf{v} \rangle \quad (1.18)$$

$$= \langle \mathbf{A} \mathbf{v}, \mathbf{v} \rangle \quad (1.19)$$

$$= \langle \mathbf{v}, \mathbf{A}^T \mathbf{v} \rangle \quad (1.20)$$

$$= \langle \mathbf{v}, \mathbf{A} \mathbf{v} \rangle \quad (1.21)$$

$$= \langle \mathbf{v}, \lambda \mathbf{v} \rangle \quad (1.22)$$

$$= \overline{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle \quad (1.23)$$

Since $\mathbf{v} \neq 0$, we have $\langle \mathbf{v}, \mathbf{v} \rangle \neq 0$. Thus $\lambda = \overline{\lambda}$, which means $\lambda \in \mathbb{R}$. \square

Corollary 1.4.3

There is an orthogonal eigenbasis for a symmetric matrix \mathbf{A} if all the eigenvalues of \mathbf{A} are all different.

Proof. Assume that the symmetric matrix \mathbf{A} has different real eigenvalues λ and μ with corresponding eigenvectors \mathbf{v} and \mathbf{w} , that means,

$$\mathbf{A} \mathbf{v} = \lambda \mathbf{v} \text{ and } \mathbf{A} \mathbf{w} = \mu \mathbf{w}$$

The relation

$$\lambda(\mathbf{v}, \mathbf{w}) = (\lambda \mathbf{v}, \mathbf{w}) \quad (1.24)$$

$$= (\mathbf{A} \mathbf{v}, \mathbf{w}) \quad (1.25)$$

$$= (\mathbf{v}, \mathbf{A}^T \mathbf{w}) \quad (1.26)$$

$$= (\mathbf{v}, \mathbf{A} \mathbf{w}) \quad (1.27)$$

$$= (\mathbf{v}, \mu \mathbf{w}) \quad (1.28)$$

$$= \overline{\mu}(\mathbf{v}, \mathbf{w}) \quad (1.29)$$

$$= \mu(\mathbf{v}, \mathbf{w}) \quad (1.30)$$

is only possible if $(\mathbf{v}, \mathbf{w}) = 0$ if $\lambda \neq \mu$

Therefore, the eigenvectors \mathbf{v} and \mathbf{w} are orthogonal. \square

Theorem 1.4.4: Spectral theorem

A symmetric matrix can be diagonalized with an orthonormal matrix S .

Proof. _____

□

need more
proof

1.4.2 Standard basis**Definition 1.4.5: Standard basis or natural basis or canonical basis**

In mathematics, the standard basis (also called natural basis or canonical basis) of a coordinate vector space (such as \mathbb{R}^n or \mathbb{C}^n) is the set of vectors, each of whose components are all zero, except one that equals 1. For example, in the case of the Euclidean plane an n -dimensional vector space \mathbb{R}^n , the standard basis vectors are a set of n vectors, each of which has a 1 in one coordinate and 0 in all other coordinates. Mathematically, the standard basis vectors can be represented as:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Here, \mathbf{e}_i denotes the i -th standard basis vector. In an n -dimensional space, the set of standard basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ forms a basis for the space. This means that any vector \mathbf{v} in the space can be expressed as a linear combination of these standard basis vectors:

$$\mathbf{v} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_n \mathbf{e}_n$$

where a_1, a_2, \dots, a_n are the coordinates of the vector \mathbf{v} with respect to the standard basis.

1.4.3 Diagonal matrix

Definition 1.4.6: Diagonal matrix

A diagonal matrix is a matrix in which all off-diagonal entries are zero. That is, the matrix $\mathbf{D} = (d_{ij})$ with n columns and n rows is diagonal if

$$\forall i, j \in 1, 2, \dots, n, i \neq j \implies d_{ij} = 0$$

However, the main diagonal entries are unrestricted.

More often, however, diagonal matrix refers to square matrices, which can be specified explicitly as a square diagonal matrix. A square diagonal matrix is a symmetric matrix, so this can also be called a symmetric diagonal matrix. The main diagonal is the line of entries running from the top left to the bottom right of the matrix. Mathematically, a diagonal matrix $\mathbf{D}_{n \times n}$ can be represented as:

$$\mathbf{D}_{n \times n} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

where d_1, d_2, \dots, d_n are the diagonal elements of the matrix.

Corollary 1.4.7

A diagonal matrix is symmetric.

$$\mathbf{D}^T = \mathbf{D}$$

Corollary 1.4.8

A diagonal matrix $\mathbf{D}_{n \times n} = \text{diag}(d_1, \dots, d_n)$ is invertible if and only if all its diagonal elements d_1, \dots, d_n are non-zero. The inverse of a diagonal matrix is also a diagonal matrix, with each diagonal element being the reciprocal of the original diagonal element.

$$\mathbf{D}_{n \times n}^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}$$

where $d_i \neq 0 \quad \forall i \in \{1, 2, \dots, n\}$

$$\text{diag}(d_1, \dots, d_n)^{-1} = \text{diag}(d_1^{-1}, \dots, d_n^{-1})$$

Corollary 1.4.9

The eigenvalues of a diagonal matrix $\mathbf{D}_{n \times n}$ are its diagonal elements d_1, d_2, \dots, d_n , and the eigenvectors are the standard basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ (1.4.2).

Corollary 1.4.10

The product of two diagonal matrices $\mathbf{D}_{n \times n}$ and $\mathbf{D}'_{n \times n}$ is another diagonal matrix, with diagonal elements being the product of the corresponding diagonal elements of the original matrices.

$$\mathbf{D}_{n \times n} \mathbf{D}'_{n \times n} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} d'_1 & 0 & \cdots & 0 \\ 0 & d'_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d'_n \end{bmatrix} = \begin{bmatrix} d_1 d'_1 & 0 & \cdots & 0 \\ 0 & d_2 d'_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n d'_n \end{bmatrix}$$

$$\text{diag}(a_1, \dots, a_n) \text{diag}(b_1, \dots, b_n) = \text{diag}(a_1 b_1, \dots, a_n b_n)$$

Corollary 1.4.11

For matrix multiplication:

$$\text{diag}(a_1, \dots, a_n) + \text{diag}(b_1, \dots, b_n) = \text{diag}(a_1 + b_1, \dots, a_n + b_n)$$

Corollary 1.4.12

The trace of a diagonal matrix $\mathbf{D}_{n \times n}$ (the sum of its diagonal elements) is equal to the sum of its eigenvalues.

$$\text{Tr}(\mathbf{D}_{n \times n}) = \sum_{i=1}^n d_i$$

Corollary 1.4.13

The determinant of a diagonal matrix $\mathbf{D}_{n \times n}$ is the product of its diagonal elements.

$$\det(\mathbf{D}_{n \times n}) = \prod_{i=1}^n d_i$$

Corollary 1.4.14

For a diagonal matrix $\mathbf{D}_{n \times n}$ with diagonal elements d_1, d_2, \dots, d_n , the k -th power of $\mathbf{D}_{n \times n}$ is another diagonal matrix, which can be expressed as:

- When $k = 0$, $\mathbf{D}_{n \times n}^0$ is the identity matrix \mathbf{I} , provided all diagonal elements d_i are non-zero ($d_i \neq 0$):

$$\mathbf{D}_{n \times n}^0 = \mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- For positive integers k , $\mathbf{D}_{n \times n}^k$ is computed by raising each diagonal element to the k -th power:

$$\mathbf{D}_{n \times n}^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}$$

- For negative integers $-k$, $\mathbf{D}_{n \times n}^{-k}$ is computed by taking the reciprocal of each diagonal element raised to the k -th power, assuming all diagonal elements d_i are non-zero ($d_i \neq 0$):

$$\mathbf{D}_{n \times n}^{-k} = \begin{bmatrix} d_1^{-k} & 0 & \cdots & 0 \\ 0 & d_2^{-k} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^{-k} \end{bmatrix} = \begin{bmatrix} \frac{1}{d_1^k} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2^k} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n^k} \end{bmatrix}$$

- When k is a fraction, the power of a diagonal matrix $\mathbf{D}_{n \times n}$ is still well-defined, provided that all the diagonal elements are non-negative ($d_i \geq 0$ if k is a real number) or positive ($d_i > 0$ if k is a complex number).

For example, if $k = \frac{1}{2}$, then $\mathbf{D}_{n \times n}^{\frac{1}{2}}$ (the square root of $\mathbf{D}_{n \times n}$) is:

$$\mathbf{D}_{n \times n}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{d_1} & 0 & \cdots & 0 \\ 0 & \sqrt{d_2} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{d_n} \end{bmatrix}$$

if $k = -\frac{1}{2}$, then $\mathbf{D}_{n \times n}^{-\frac{1}{2}}$ is:

$$\mathbf{D}_{n \times n}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{d_2}} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{d_n}} \end{bmatrix}$$

1.4.4 Degree matrix

Definition 1.4.15: Degree matrix

Given a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with $|\mathbf{V}| = n$, \mathbf{V} is the vertex set, and \mathbf{E} is the edge set, the Degree matrix $\mathbf{D}_{n \times n}$ for \mathbf{G} is a $n \times n$ diagonal matrix defined as

$$\mathbf{D}_{ij} = \begin{cases} \deg(v_i) & i = j \\ 0 & i \neq j \end{cases}$$

where the degree $\deg(v_i)$ of a vertex counts the number of times an edge terminates at that vertex.

In an **undirected graph**, this means that each loop increases the degree of a vertex by two.

In a **directed graph**, the term degree may refer either to indegree (the number of incoming edges at each vertex) or outdegree (the number of outgoing edges at each vertex)

Corollary 1.4.16

Since the Degree matrix $\mathbf{D}_{n \times n}$ for \mathbf{G} is a $n \times n$ diagonal matrix, it is also symmetric, meaning.

$$\mathbf{D}_{ij} = \mathbf{D}_{ji}$$

1.4.5 Adjacency matrix

Definition 1.4.17: Adjacency matrix

Given a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with $|\mathbf{V}| = n$, \mathbf{V} is the vertex set, and \mathbf{E} is the edge set, the Adjacency matrix $\mathbf{A}_{n \times n}$ for \mathbf{G} is a $n \times n$ matrix defined as

$$\mathbf{A}_{ij} = \begin{cases} 1 & (v_i, v_j) \in E \\ 0 & (v_i, v_j) \notin E \end{cases}$$

where (v_i, v_j) is an edge in \mathbf{E} .

A matrix with rows and columns labeled by nodes, whose i -th row, j -th column component a_{ij} is 1 if node i is a neighbor of node j , or 0 otherwise.

Corollary 1.4.18

The diagonal elements of the Adjacency matrix are all zero, since edges from a vertex to itself (loops) are not allowed in simple graphs.

$$\mathbf{A}_{ii} = 0$$

Corollary 1.4.19

For an undirected graph, the Adjacency matrix is symmetric, meaning.

$$\mathbf{A}_{ij} = \mathbf{A}_{ji}$$

For a directed graph, the Adjacency matrix may not be symmetric.

Definition 1.4.20: Adjacency list

Adjacency list \mathbf{A} list of lists of nodes whose i -th component is the list of node i 's neighbors.

1.4.6 Weighted Adjacency matrix**Definition 1.4.21: Weighted Adjacency matrix**

Let $\mathbf{G} = (\mathbf{V}, \mathbf{W})$ be an undirected graph, where \mathbf{V} is the n vertex set, and $\mathbf{W}_{n \times n}$ is the $n \times n$ Weighted Adjacency matrix. Let w_{ij} be the non-negative weight between nodes i, j .

$$\mathbf{W}_{ij} = \begin{cases} w_{ij} & (v_i, v_j) \in E \\ 0 & (v_i, v_j) \notin E \end{cases}$$

An edge exists between nodes i and j if $w_{ij} > 0$.

For an unweighted graph, $w_{ij} = 1$ if there is an edge; $w_{ij} = 0$ otherwise.

Corollary 1.4.22

For undirected graphs, the Weighted Adjacency matrix is symmetric, meaning

$$\mathbf{W}_{ij} = \mathbf{W}_{ji}$$

For a directed graph, the Weighted Adjacency matrix may not be symmetric.

Corollary 1.4.23

For a simple graph, the diagonal elements of the Weighted Adjacency matrix are all 0, indicating no self-loops.

$$\mathbf{W}_{ii} = 0 \quad \forall i \in \{0, 1, \dots, n\}$$

Corollary 1.4.24

The diagonal elements of the Degree Matrix \mathbf{D} are the sums of the elements in each row of the Weighted Adjacency Matrix \mathbf{W} .

Let the diagonal element \mathbf{D}_{ii} represents the degree of vertex i , which is the sum of the weights of the edges connected to vertex i . Specifically:

$$\mathbf{D}_{ii} = \sum_{j=0}^{n-1} \mathbf{W}_{ij}$$

Therefore, the Degree Matrix \mathbf{D} can be expressed as:

$$\mathbf{D} = \text{diag} \left(\sum_{j=0}^{n-1} \mathbf{W}_{ij} \right)$$

where $\text{diag}(\cdot)$ denotes constructing a diagonal matrix with the values inside the parentheses as its diagonal elements.

1.4.7 Weighted vertex-edge matrix**Definition 1.4.25: Weighted vertex-edge matrix**

The edge between nodes i and j can be represented as an n -dimensional vector:

$$\mathbf{e} = \sqrt{w_{ij}} [0, \dots, 0, 1, 0, \dots, 0, -1, 0, \dots, 0]^T$$

where the i -th location is 1, the j -th location is -1, and other locations are 0.

$$\mathbf{e}(\tau) = \begin{cases} w_{ij}^{\frac{1}{2}} & \text{if } \tau \text{ is } i, \\ -w_{ij}^{\frac{1}{2}} & \text{if } \tau \text{ is } j, \\ 0 & \text{otherwise} \end{cases}$$

Here, $\mathbf{e}(\tau)$ denotes the τ -th value of the edge vector \mathbf{e} . For undirected graphs, the negative endpoint can be either at the $\tau = i$ location or at the $\tau = j$ location. Let $\mathbf{E}_{n \times m} = [\mathbf{e}_1, \dots, \mathbf{e}_m]$ be the weighted vertex-edge matrix of size $n \times m$ obtained by taking the m edge vectors as its columns. (The order of the edges is arbitrary.)

1.4.8 Normalized weighted vertex-edge matrix

Definition 1.4.26: Normalized weighted vertex-edge matrix

Let \mathbf{D} be the Degree matrix of the graph with the i -th diagonal element:

$$d_i = \sum_{j=1}^n w_{ij}$$

Let $\mathcal{E}_{n \times m} = [\epsilon_1, \dots, \epsilon_m]$ be the Normalized weighted vertex-edge matrix with each n -dimensional edge ϵ vector normalized:

$$\mathcal{E}_{n \times m} = \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{E}_{n \times m}$$

with $\epsilon = \sqrt{w_{ij}} \left[0, \dots, 0, \frac{1}{\sqrt{d_i}}, 0, \dots, 0, -\frac{1}{\sqrt{d_j}}, 0, \dots, 0 \right]^T$ where the i -th location is $\sqrt{d_i}$, the j -th location is $-\sqrt{d_j}$, and other locations are 0.

$$\epsilon(\tau) = \begin{cases} d_i^{-\frac{1}{2}} w_{ij}^{\frac{1}{2}} & \text{if } \tau \text{ is } i, \\ -d_j^{-\frac{1}{2}} w_{ij}^{\frac{1}{2}} & \text{if } \tau \text{ is } j, \\ 0 & \text{otherwise} \end{cases}$$

Here, $\epsilon(\tau)$ denotes the τ -th value of the edge vector ϵ .

$$\epsilon(\tau) = \mathbf{D}^{-\frac{1}{2}} \mathbf{e}(\tau)$$

Proof.

$$\epsilon(\tau) = \mathbf{D}^{-\frac{1}{2}} \mathbf{e}(\tau) \quad (1.31)$$

$$= \begin{bmatrix} d_1^{-\frac{1}{2}} & 0 & \cdots & 0 \\ 0 & d_2^{-\frac{1}{2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & d_n^{-\frac{1}{2}} \end{bmatrix} \sqrt{w_{ij}} \begin{bmatrix} 0 \\ \cdots \\ 0 \\ 1 \\ 0 \\ \cdots \\ 0 \\ -1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \quad (1.32)$$

$$= \sqrt{w_{ij}} \begin{bmatrix} 0 \\ \cdots \\ 0 \\ \frac{1}{\sqrt{d_i}} \\ 0 \\ \cdots \\ 0 \\ -\frac{1}{\sqrt{d_j}} \\ 0 \\ \cdots \\ 0 \end{bmatrix} \quad (1.33)$$

$$= \sqrt{w_{ij}} \left[0, \cdots, 0, \frac{1}{\sqrt{d_i}}, 0, \cdots, 0, -\frac{1}{\sqrt{d_j}}, 0, \cdots, 0 \right]^{\mathbf{T}} \quad (1.34)$$

□

1.4.9 Laplacian matrix

Definition 1.4.27: Laplacian matrix

Given a simple \mathbf{G} with n vertices v_1, \dots, v_n , its Laplacian matrix $\mathbf{L}_{n \times n}$ is defined element-wise as

$$\mathbf{L}_{ij} = \begin{cases} \deg(v_i) & i = j \\ -w_{ij} & i \neq j \text{ and } (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{cases}$$

or equivalently by the matrix

$$\mathbf{L}_{n \times n} = \mathbf{D}_{n \times n} - \mathbf{W}_{n \times n}$$

where $\mathbf{D}_{n \times n}$ is the Degree matrix and $\mathbf{W}_{n \times n}$ is the Weighted Adjacency matrix of the graph \mathbf{G} .

For an unweighted graph, $w_{ij} = 1$ if there is an edge. then

$$\mathbf{L}_{ij} = \begin{cases} \deg(v_i) & i = j \\ -1 & i \neq j \text{ and } (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{cases}$$

or equivalently by the matrix

$$\mathbf{L}_{n \times n} = \mathbf{D}_{n \times n} - \mathbf{A}_{n \times n}$$

where $\mathbf{D}_{n \times n}$ is the Degree matrix and $\mathbf{A}_{n \times n}$ is the Adjacency matrix of the graph \mathbf{G} .

Since \mathbf{G} is a simple graph, \mathbf{A} only contains 1s or 0s and its diagonal elements are all 0s.

Corollary 1.4.28

For an undirected graph, the Laplacian matrix $\mathbf{L}_{n \times n}$ is a symmetric matrix.

Proof. The Laplacian matrix $\mathbf{L}_{n \times n}$ is defined as $\mathbf{L}_{n \times n} = \mathbf{D}_{n \times n} - \mathbf{W}_{n \times n}$, where $\mathbf{D}_{n \times n}$ is the degree matrix and $\mathbf{W}_{n \times n}$ is the weighted adjacency matrix.

For an undirected graph, the adjacency matrix $\mathbf{W}_{n \times n}$ is symmetric, meaning $\mathbf{W}_{n \times n} = \mathbf{W}_{n \times n}^T$.

The degree matrix $\mathbf{D}_{n \times n}$ is diagonal and therefore symmetric.

Since both $\mathbf{D}_{n \times n}$ and $\mathbf{W}_{n \times n}$ are symmetric, their difference $\mathbf{L}_{n \times n} = \mathbf{D}_{n \times n} - \mathbf{W}_{n \times n}$ is also symmetric. \square

Corollary 1.4.29

For an undirected graph, the eigenvalues of the Laplacian matrix \mathbf{L} are real and non-negative.

Proof. Since the Laplacian matrix \mathbf{L} is symmetric, then its eigenvalues are real (1.4.1). \square

Corollary 1.4.30

For the undirected graph that both the Adjacency matrix \mathbf{A} and the Laplacian matrix \mathbf{L} are symmetric, and that row- and column-sums of the Laplacian matrix are all zeros (which directly implies that the Laplacian matrix is singular).

Corollary 1.4.31

Let $\mathbf{E}_{n \times m} = [\mathbf{e}_1, \dots, \mathbf{e}_m]$ be the $n \times m$ weighted vertex-edge matrix in the an undirected graph of n vertices and m edges (1.4.7).

The graph Laplacian matrix can be computed as follows

$$\mathbf{L}_{n \times n} = \mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T = \sum_{t=1}^m \mathbf{e}_t \mathbf{e}_t^T$$

Proof. Since $\mathbf{E}_{n \times m} = [\mathbf{e}_1, \dots, \mathbf{e}_m]$ be the $n \times m$ weighted vertex-edge matrix in the an undirected graph of n vertices and m edges.

Let $\mathbf{e}(\tau)$ denotes the τ -th value of the edge vector \mathbf{e} . For undirected graphs, the negative endpoint can be either at the $\tau = i$ location or at the $\tau = j$ location.

$$\mathbf{e}(\tau) = \begin{cases} w_{ij}^{\frac{1}{2}} & \text{if } \tau \text{ is } i, \\ -w_{ij}^{\frac{1}{2}} & \text{if } \tau \text{ is } j, \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_m \end{bmatrix} \begin{bmatrix} \mathbf{e}_1^T \\ \vdots \\ \mathbf{e}_m^T \end{bmatrix} \quad (1.35)$$

$$= \sum_{t=1}^m \mathbf{e}_t \mathbf{e}_t^T \quad (1.36)$$

$$= \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & w_{ij} & \cdots & -w_{ij} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & -w_{ij} & \cdots & w_{ij} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \quad (1.37)$$

For each \mathbf{e}_t , its non-zero elements are at positions i and j , with values $\sqrt{w_{ij}}$ and $-\sqrt{w_{ij}}$ respectively. Therefore, $\mathbf{e}_t \mathbf{e}_t^T$ is an $n \times n$ matrix with w_{ij} at positions (i, i) and (j, j) , $-w_{ij}$ at positions (i, j) and (j, i) , and 0 elsewhere.

Summing all $\mathbf{e}_t \mathbf{e}_t^T$ matrices results in a matrix $\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T$ where the diagonal elements are the degrees of the vertices (i.e., the degree matrix $\mathbf{D}_{n \times n}$), and the off-diagonal elements are the negative weights of the edges (i.e., the negative of the weighted adjacency matrix $\mathbf{W}_{n \times n}$).

Specifically, for the i -th row and i -th column element of the matrix $\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T$:

$$(\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T)_{ii} = \sum_{t=1}^m (\mathbf{e}_t \mathbf{e}_t^T)_{ii} = \sum_{j \neq i} w_{ij} = d_i$$

where d_i is the degree of vertex i (1.4.4).

For the i -th row and j -th column element of the matrix $\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T$ (where $i \neq j$):

$$(\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T)_{ij} = \sum_{t=1}^m (\mathbf{e}_t \mathbf{e}_t^T)_{ij} = -w_{ij}$$

where w_{ij} is the weight of the edge between vertices i and j (1.4.6).

Thus, $\mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T = \mathbf{D}_{n \times n} - \mathbf{W}_{n \times n} = \mathbf{L}_{n \times n}$. □

1.4.10 Normalized Laplacian matrix

Definition 1.4.32: Symmetrically normalized Laplacian matrix

The symmetrically normalized Laplacian matrix can be computed as:

$$\mathcal{L}_{sym} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}}$$

The elements of \mathcal{L}_{sym} are thus given by

$$(\mathcal{L}_{sym})_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } \deg(v_i) \neq 0 \\ -\frac{w_{ij}}{\sqrt{\deg(v_i) \deg(v_j)}} & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise} \end{cases}$$

Proof. The Degree matrix $\mathbf{D}_{n \times n}$ is a diagonal matrix, and when $d_i \neq 0 \ \forall i \in \{1, 2, \dots, n\}$, then

$$\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{D}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{d_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{d_n}} \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{d_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{d_n}} \end{bmatrix} \quad (1.38)$$

$$= \begin{bmatrix} \frac{d_1}{\sqrt{d_1} \sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{d_2}{\sqrt{d_2} \sqrt{d_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{d_n}{\sqrt{d_n} \sqrt{d_n}} \end{bmatrix} \quad (1.39)$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (1.40)$$

$$= \mathbf{I}_{n \times n} \quad (1.41)$$

Since $\mathbf{L}_{n \times n} = \mathbf{D}_{n \times n} - \mathbf{W}_{n \times n}$, then

$$\mathcal{L}_{sym} = \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{L}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.42)$$

$$= \mathbf{D}_{n \times n}^{-\frac{1}{2}} (\mathbf{D}_{n \times n} - \mathbf{W}_{n \times n}) \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.43)$$

$$= \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{D}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} - \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{W}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.44)$$

$$= \mathbf{I}_{n \times n} - \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{W}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.45)$$

And we have

$$\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{W}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{d_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{d_n}} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{d_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{d_n}} \end{bmatrix} \quad (1.46)$$

$$= \begin{bmatrix} \frac{w_{11}}{\sqrt{d_1}} & \frac{w_{12}}{\sqrt{d_1}} & \cdots & \frac{w_{1n}}{\sqrt{d_1}} \\ \frac{w_{21}}{\sqrt{d_2}} & \frac{w_{22}}{\sqrt{d_2}} & \cdots & \frac{w_{2n}}{\sqrt{d_2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{\sqrt{d_n}} & \frac{w_{n2}}{\sqrt{d_n}} & \cdots & \frac{w_{nn}}{\sqrt{d_n}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{d_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{d_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{d_n}} \end{bmatrix} \quad (1.47)$$

$$= \begin{bmatrix} \frac{w_{11}}{\sqrt{d_1}\sqrt{d_1}} & \frac{w_{12}}{\sqrt{d_1}\sqrt{d_2}} & \cdots & \frac{w_{1n}}{\sqrt{d_1}\sqrt{d_n}} \\ \frac{w_{21}}{\sqrt{d_2}\sqrt{d_1}} & \frac{w_{22}}{\sqrt{d_2}\sqrt{d_2}} & \cdots & \frac{w_{2n}}{\sqrt{d_2}\sqrt{d_n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{\sqrt{d_n}\sqrt{d_1}} & \frac{w_{n2}}{\sqrt{d_n}\sqrt{d_2}} & \cdots & \frac{w_{nn}}{\sqrt{d_n}\sqrt{d_n}} \end{bmatrix} \quad (1.48)$$

We know that the diagonal elements of the Weighted Adjacency matrix are all zero, that means $w_{ii} = 0 \quad \forall i \in \{0, 1, \dots, n\}$ (1.4.6). And the $w_{ij} = 0$ if $i \neq j$ and vertex i is not adjacent to vertex j , then

$$\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{W}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} = \begin{bmatrix} 0 & \frac{w_{12}}{\sqrt{d_1}\sqrt{d_2}} & \cdots & \frac{w_{1n}}{\sqrt{d_1}\sqrt{d_n}} \\ \frac{w_{21}}{\sqrt{d_2}\sqrt{d_1}} & 0 & \cdots & \frac{w_{2n}}{\sqrt{d_2}\sqrt{d_n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{\sqrt{d_n}\sqrt{d_1}} & \frac{w_{n2}}{\sqrt{d_n}\sqrt{d_2}} & \cdots & 0 \end{bmatrix}$$

Thus, we have

$$\mathcal{L}_{sym} = \mathbf{I}_{n \times n} - \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{W}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.49)$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} - \begin{bmatrix} 0 & \frac{w_{12}}{\sqrt{d_1}\sqrt{d_2}} & \cdots & \frac{w_{1n}}{\sqrt{d_1}\sqrt{d_n}} \\ \frac{w_{21}}{\sqrt{d_2}\sqrt{d_1}} & 0 & \cdots & \frac{w_{2n}}{\sqrt{d_2}\sqrt{d_n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{\sqrt{d_n}\sqrt{d_1}} & \frac{w_{n2}}{\sqrt{d_n}\sqrt{d_2}} & \cdots & 0 \end{bmatrix} \quad (1.50)$$

$$= \begin{bmatrix} 1 & -\frac{w_{12}}{\sqrt{d_1}\sqrt{d_2}} & \cdots & -\frac{w_{1n}}{\sqrt{d_1}\sqrt{d_n}} \\ -\frac{w_{21}}{\sqrt{d_2}\sqrt{d_1}} & 1 & \cdots & -\frac{w_{2n}}{\sqrt{d_2}\sqrt{d_n}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{w_{n1}}{\sqrt{d_n}\sqrt{d_1}} & -\frac{w_{n2}}{\sqrt{d_n}\sqrt{d_2}} & \cdots & 1 \end{bmatrix} \quad (1.51)$$

We know that the degree of the vertex i is $d_i = \deg(v_i)$, then the elements of \mathcal{L}_{sym} are

$$(\mathcal{L}_{sym})_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } \deg(v_i) \neq 0 \\ -\frac{w_{ij}}{\sqrt{\deg(v_i)\deg(v_j)}} & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise} \end{cases}$$

□

Corollary 1.4.33

$$\mathcal{L}_{sym} = \mathcal{E}\mathcal{E}^T = \sum_{t=1}^m \epsilon_t \epsilon_t^T$$

Proof. Since $\mathbf{L}_{n \times n} = \mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T = \sum_{t=1}^m \mathbf{e}_t \mathbf{e}_t^T$, $\mathcal{E}_{n \times m} = \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{E}_{n \times m}$ and $\mathbf{D}_{n \times n}^T = \mathbf{D}_{n \times n}$, then

$$\mathcal{L}_{sym} = \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{L}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.52)$$

$$= \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T (\mathbf{D}_{n \times n}^T)^{-\frac{1}{2}} \quad (1.53)$$

$$= \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{E}_{n \times m} \mathbf{E}_{n \times m}^T (\mathbf{D}_{n \times n}^{-\frac{1}{2}})^T \quad (1.54)$$

$$= (\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{E}_{n \times m}) (\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{E}_{n \times m})^T \quad (1.55)$$

$$= \mathcal{E} \mathcal{E}^T \quad (1.56)$$

And we have $\epsilon_t = \mathbf{D}^{-\frac{1}{2}} \mathbf{e}_t$, then

$$\mathcal{L}_{sym} = \mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{L}_{n \times n} \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.57)$$

$$= \mathbf{D}_{n \times n}^{-\frac{1}{2}} \left(\sum_{t=1}^m \mathbf{e}_t \mathbf{e}_t^T \right) \mathbf{D}_{n \times n}^{-\frac{1}{2}} \quad (1.58)$$

$$= \sum_{t=1}^m \left(\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{e}_t \mathbf{e}_t^T \mathbf{D}_{n \times n}^{-\frac{1}{2}} \right) \quad (1.59)$$

$$= \sum_{t=1}^m \left[\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{e}_t \mathbf{e}_t^T (\mathbf{D}_{n \times n}^T)^{-\frac{1}{2}} \right] \quad (1.60)$$

$$= \sum_{t=1}^m \left[\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{e}_t \mathbf{e}_t^T (\mathbf{D}_{n \times n}^{-\frac{1}{2}})^T \right] \quad (1.61)$$

$$= \sum_{t=1}^m \left(\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{e}_t \right) \left(\mathbf{D}_{n \times n}^{-\frac{1}{2}} \mathbf{e}_t \right)^T \quad (1.62)$$

$$= \sum_{t=1}^m \epsilon_t \epsilon_t^T \quad (1.63)$$

□

1.4.11 Eigendecomposition

Definition 1.4.34: Eigendecomposition

Let \mathbf{A} be a square $n \times n$ matrix with n linearly independent eigenvectors q_i (where $i = 1, \dots, n$). Then \mathbf{A} can be factored as

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$$

where \mathbf{Q} is the square $n \times n$ matrix whose i -th column is the eigenvector q_i of \mathbf{A} , and $\mathbf{\Lambda}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues, $\Lambda_{ii} = \lambda_i$. Note that only diagonalizable matrices can be factorized in this way.

The n eigenvectors q_i are usually normalized, but they don't have to be. A non-normalized set of n eigenvectors, v_i can also be used as the columns of \mathbf{Q} .

Fact 1.4.35

Assume \mathbf{A} has non-degenerate eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ and corresponding linearly independent eigen-

vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$ which can be denoted $\begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1k} \end{bmatrix}, \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2k} \end{bmatrix}, \dots, \begin{bmatrix} x_{k1} \\ x_{k2} \\ \vdots \\ x_{kk} \end{bmatrix}$.

Define the matrices composed of eigenvectors

$$\mathbf{Q} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_k \end{bmatrix} \quad (1.64)$$

$$= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{kk} \end{bmatrix} \quad (1.65)$$

and eigenvalues

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}.$$

where Λ is a diagonal matrix. Then

$$\mathbf{A}\mathbf{Q} = \mathbf{A} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_k \end{bmatrix} \quad (1.66)$$

$$= \begin{bmatrix} \mathbf{A}\mathbf{X}_1 & \mathbf{A}\mathbf{X}_2 & \cdots & \mathbf{A}\mathbf{X}_k \end{bmatrix} \quad (1.67)$$

$$= \begin{bmatrix} \lambda_1 \mathbf{X}_1 & \lambda_2 \mathbf{X}_2 & \cdots & \lambda_k \mathbf{X}_k \end{bmatrix} \quad (1.68)$$

$$= \begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{21} & \cdots & \lambda_k x_{k1} \\ \lambda_1 x_{12} & \lambda_2 x_{22} & \cdots & \lambda_k x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 x_{1k} & \lambda_2 x_{2k} & \cdots & \lambda_k x_{kk} \end{bmatrix} \quad (1.69)$$

$$= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{kk} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix} \quad (1.70)$$

$$= \mathbf{Q}\Lambda \quad (1.71)$$

giving the amazing decomposition of \mathbf{A} into a similarity transformation involving \mathbf{Q} and Λ ,

$$\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^{-1}$$

The fact that this decomposition is always possible for a square matrix \mathbf{A} as long as \mathbf{Q} is a square matrix is known in this work as the Eigendecomposition theorem.

Claim: How to Perform Eigendecomposition?

To perform Eigendecomposition on a matrix, follow these steps:

1. Find the Eigenvalues:

Solve the characteristic equation:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

Here, \mathbf{A} is the square matrix, λ is the eigenvalue, and \mathbf{I} is the identity matrix of the same dimension as \mathbf{A} .

2. Find the Eigenvectors:

For each eigenvalue λ , substitute it back into the equation:

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = 0$$

This represents a system of linear equations where \mathbf{v} is the eigenvector corresponding to the eigenvalue λ .

3. Construct the Eigenvector Matrix \mathbf{Q} :

Place all the eigenvectors as columns in the matrix \mathbf{Q} . If there are n distinct eigenvalues, \mathbf{Q} will be an $n \times n$ matrix.

4. Form the Diagonal Matrix Λ :

Construct a diagonal matrix Λ by placing the eigenvalues on its diagonal

5. Calculate the Inverse of \mathbf{Q} :

Find \mathbf{Q}^{-1} , the inverse of the eigenvector matrix \mathbf{Q} , if the matrix is invertible.

6. Combine to Form the Decomposition:

Express the original matrix \mathbf{A} as

$$\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^{-1}$$

Corollary 1.4.36

For general positive integer powers $n > 0$,

$$\mathbf{A}^n = \mathbf{Q}\Lambda^n\mathbf{Q}^{-1}$$

Corollary 1.4.37

The inverse of \mathbf{A} is

$$\mathbf{A}^{-1} = (\mathbf{Q}\Lambda\mathbf{Q}^{-1})^{-1} \tag{1.72}$$

$$= (\mathbf{Q}^{-1})^{-1}(\Lambda)^{-1}(\mathbf{Q})^{-1} \tag{1.73}$$

$$= \mathbf{Q}\Lambda^{-1}\mathbf{Q}^{-1} \tag{1.74}$$

Corollary 1.4.38

The inverse of the diagonal matrix Λ is trivially given by

$$\Lambda^{-1} = \begin{bmatrix} \lambda_1^{-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^{-1} \end{bmatrix}$$

1.4.12 Matrix Exponential**Definition 1.4.39: Matrix Exponential**

The power series that defines the exponential map e^x also defines a map between matrices. In particular,

$$\exp(\mathbf{A}) = e^{\mathbf{A}} \tag{1.75}$$

$$= \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!} \tag{1.76}$$

$$= \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}\mathbf{A}}{2!} + \frac{\mathbf{A}\mathbf{A}\mathbf{A}}{3!} + \cdots \tag{1.77}$$

converges for any square matrix \mathbf{A} , where \mathbf{I} is the identity matrix.

Corollary 1.4.40

Since Λ is a diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}$$

Then

$$\exp(\Lambda) = e^\Lambda \tag{1.78}$$

$$= \sum_{n=0}^{\infty} \frac{\Lambda^n}{n!} \tag{1.79}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^n \end{bmatrix} \tag{1.80}$$

$$= \begin{bmatrix} \sum_{n=0}^{\infty} \frac{\lambda_1^n}{n!} & 0 & \cdots & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{\lambda_2^n}{n!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{n=0}^{\infty} \frac{\lambda_k^n}{n!} \end{bmatrix} \tag{1.81}$$

$$= \begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_k} \end{bmatrix} \tag{1.82}$$

Since $\mathbf{A}^n = \mathbf{Q}\Lambda^n\mathbf{Q}^{-1}$, then

$$\exp(\mathbf{A}) = e^{\mathbf{A}} \tag{1.83}$$

$$= \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!} \tag{1.84}$$

$$= \sum_{n=0}^{\infty} \frac{\mathbf{Q}\Lambda^n\mathbf{Q}^{-1}}{n!} \tag{1.85}$$

$$= \mathbf{Q} \left(\sum_{n=0}^{\infty} \frac{\Lambda^n}{n!} \right) \mathbf{Q}^{-1} \tag{1.86}$$

$$= \mathbf{Q} e^\Lambda \mathbf{Q}^{-1} \tag{1.87}$$

so $e^{\mathbf{A}}$ can be found using Λ .

Corollary 1.4.41

For the zero matrix $\mathbf{A} = \mathbf{0}$,

$$\exp(\mathbf{A}) = e^{\mathbf{0}} = \mathbf{I}$$

i.e., the identity matrix.

In general,

$$e^{\mathbf{A}}e^{-\mathbf{A}} = e^{\mathbf{0}} = \mathbf{I}$$

so the exponential of a matrix is always invertible, with inverse the exponential of the negative of the matrix.

Corollary 1.4.42

$$\exp(\mathbf{A}^T) = \exp(\mathbf{A})^T$$

Corollary 1.4.43

If \mathbf{Q} is invertible, then

$$e^{\mathbf{Q}\mathbf{A}\mathbf{Q}^{-1}} = \mathbf{Q}e^{\mathbf{A}}\mathbf{Q}^{-1}$$

Corollary 1.4.44

If $\mathbf{AB} = \mathbf{BA}$, then

$$e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}}$$

Corollary 1.4.45

$$e^{a\mathbf{A}}e^{b\mathbf{A}} = e^{(a+b)\mathbf{A}}$$

1.4.13 QR decomposition**Definition 1.4.46: QR decomposition or QR factorization or QU factorization**

Any real square matrix \mathbf{A} may be decomposed as

$$\mathbf{A} = \mathbf{QR}$$

where \mathbf{Q} is an orthogonal matrix (its columns are orthogonal unit vectors meaning

$$\mathbf{Q}^T = \mathbf{Q}^{-1} \quad \text{or} \quad \mathbf{Q}^T\mathbf{Q} = \mathbf{Q}\mathbf{Q}^T = \mathbf{I}$$

and \mathbf{R} is an upper triangular matrix (also called right triangular matrix). If \mathbf{A} is invertible, then the factorization is unique if we require the diagonal elements of \mathbf{R} to be positive.

If instead \mathbf{A} is a complex square matrix, then there is a decomposition $\mathbf{A} = \mathbf{QR}$ where \mathbf{Q} is a unitary matrix (so the conjugate transpose $\mathbf{Q}^\dagger = \mathbf{Q}^{-1}$).

Corollary 1.4.47

QR decomposition is often used to solve linear systems, eigenvalue problems and the linear least squares (LLS) problem and is the basis for a particular eigenvalue algorithm, the QR algorithm.

Corollary 1.4.48

If \mathbf{A} has n linearly independent columns, then the first n columns of \mathbf{Q} form an orthonormal basis for the column space of \mathbf{A} . More generally, the first k columns of \mathbf{Q} form an orthonormal basis for the span of the first k columns of \mathbf{A} for any $1 \leq k \leq n$. The fact that any column k of \mathbf{A} only depends on the first k columns of \mathbf{Q} corresponds to the triangular form of \mathbf{R} .

Corollary 1.4.49

More generally, we can factor a complex $m \times n$ matrix \mathbf{A} , with $m \geq n$, as the product of an $m \times n$ unitary matrix \mathbf{Q} and an $m \times n$ upper triangular matrix \mathbf{R} . As the bottom $(m - n)$ rows of an $m \times n$ upper triangular matrix consist entirely of zeroes, it is often useful to partition \mathbf{R} , or both \mathbf{R} and \mathbf{Q} :

$$\mathbf{A} = \mathbf{QR} = \mathbf{Q} \begin{bmatrix} \mathbf{R}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ 0 \end{bmatrix} = \mathbf{Q}_1 \mathbf{R}_1$$

where \mathbf{R}_1 is an $n \times n$ upper triangular matrix, 0 is an $(m - n) \times n$ zero matrix, \mathbf{Q}_1 is $m \times n$, \mathbf{Q}_2 is $m \times (m - n)$, and \mathbf{Q}_1 and \mathbf{Q}_2 both have orthogonal columns.

Golub & Van Loan (1996) call $\mathbf{Q}_1 \mathbf{R}_1$ the **thin QR factorization** of \mathbf{A} ; Trefethen and Bau call this the **reduced QR factorization**. If \mathbf{A} is of full rank n and we require that the diagonal elements of \mathbf{R}_1 are positive then \mathbf{R}_1 and \mathbf{Q}_1 are unique, but in general \mathbf{Q}_2 is not. \mathbf{R}_1 is then equal to the upper triangular factor of the Cholesky decomposition of $\mathbf{A}^* \mathbf{A}$ ($= \mathbf{A}^T \mathbf{A}$ if \mathbf{A} is real).

Definition 1.4.50: QL, RQ and LQ decompositions

Analogously, we can define QL, RQ, and LQ decompositions, with \mathbf{L} being a lower triangular matrix.

Claim: Computing the QR decomposition: Using the Gram–Schmidt process

There are several methods for actually computing the QR decomposition, such as the **Gram–Schmidt process**, **Householder transformations**, or **Givens rotations**. Each has a number of advantages and disadvantages.

Using the Gram–Schmidt process

Consider the Gram–Schmidt process applied to the columns of the full column rank matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$$

, with inner product $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w}$ (or $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^\dagger \mathbf{w}$ for the complex case).

Define the projection:

$$\text{proj}_{\mathbf{u}} = \mathbf{a} = \frac{\langle \mathbf{u}, \mathbf{a} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}$$

then:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \\ \mathbf{u}_2 &= \mathbf{a}_2 - \text{proj}_{\mathbf{u}_1} \mathbf{a}_2, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \\ \mathbf{u}_3 &= \mathbf{a}_3 - \text{proj}_{\mathbf{u}_1} \mathbf{a}_3 - \text{proj}_{\mathbf{u}_2} \mathbf{a}_3, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}, \\ &\vdots & &\vdots \\ \mathbf{u}_k &= \mathbf{a}_k - \sum_{j=1}^{k-1} \text{proj}_{\mathbf{u}_j} \mathbf{a}_k, & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}. \end{aligned}$$

We can now express the \mathbf{a}_i s over our newly computed orthonormal basis:

$$\begin{aligned} \mathbf{a}_1 &= \langle \mathbf{e}_1, \mathbf{a}_1 \rangle \mathbf{e}_1, \\ \mathbf{a}_2 &= \langle \mathbf{e}_1, \mathbf{a}_2 \rangle \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_2 \rangle \mathbf{e}_2, \\ \mathbf{a}_3 &= \langle \mathbf{e}_1, \mathbf{a}_3 \rangle \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_3 \rangle \mathbf{e}_2 + \langle \mathbf{e}_3, \mathbf{a}_3 \rangle \mathbf{e}_3, \\ &\vdots \\ \mathbf{a}_k &= \sum_{j=1}^k \langle \mathbf{e}_j, \mathbf{a}_k \rangle \mathbf{e}_j. \end{aligned}$$

where $\langle \mathbf{e}_i, \mathbf{a}_i \rangle = \|\mathbf{u}_i\|$. This can be written in matrix form:

$$\mathbf{A} = \mathbf{Q}\mathbf{R}$$

where:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{bmatrix}$$

and

$$\mathbf{R} = \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{a}_1 \rangle & \langle \mathbf{e}_1, \mathbf{a}_2 \rangle & \langle \mathbf{e}_1, \mathbf{a}_3 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{a}_n \rangle \\ 0 & \langle \mathbf{e}_2, \mathbf{a}_2 \rangle & \langle \mathbf{e}_2, \mathbf{a}_3 \rangle & \cdots & \langle \mathbf{e}_2, \mathbf{a}_n \rangle \\ 0 & 0 & \langle \mathbf{e}_3, \mathbf{a}_3 \rangle & \cdots & \langle \mathbf{e}_3, \mathbf{a}_n \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \langle \mathbf{e}_n, \mathbf{a}_n \rangle \end{bmatrix}.$$

Claim: Relation to RQ decomposition

The RQ decomposition transforms a matrix \mathbf{A} into the product of an upper triangular matrix \mathbf{R} (also known as right-triangular) and an orthogonal matrix \mathbf{Q} . The only difference from QR decomposition is the order of these matrices.

QR decomposition is Gram–Schmidt orthogonalization of columns of \mathbf{A} , started from the first column.

RQ decomposition is Gram–Schmidt orthogonalization of rows of \mathbf{A} , started from the last row.

Claim: Advantages and disadvantages

The Gram-Schmidt process is inherently numerically unstable. While the application of the projections has an appealing geometric analogy to orthogonalization, the orthogonalization itself is prone to numerical error. A significant advantage is the ease of implementation.

Chapter 2

Classical Spectral Clustering Algorithm

2.1 Classical Spectral Clustering Algorithm

Bibliography

- [1] NICE CXone. Libretexts mathematics library, Jun 2024.