

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	Degree matrix	10
1.4.3	Adjacency matrix	10
1.4.4	Vertex-edge matrix	12
1.4.5	Laplacian matrix	12
1.4.6	Eigendecomposition	13
1.4.7	Matrix Exponential	16
2	Classical Spectral Clustering Algorithm	19
2.1	Classical Spectral Clustering Algorithm	19

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

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

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

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

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:

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

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:

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

$$= \sum_{i=1}^n \sum_{j=1}^n \overline{\mathbf{A}_{ij}} v_j \overline{w_i} \quad (1.5)$$

$$= \sum_{j=1}^n v_j \left(\sum_{i=1}^n \overline{\mathbf{A}_{ij}} \overline{w_i} \right) \quad (1.6)$$

$$= \sum_{j=1}^n v_j \left(\sum_{i=1}^n \overline{\mathbf{A}_{ij} w_i} \right) \quad (1.7)$$

$$= \sum_{j=1}^n v_j \left(\sum_{i=1}^n \overline{(\mathbf{A}^T)_{ji} w_i} \right) \quad (1.8)$$

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

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

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

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

$$= \lambda \sum_{i=1}^n v_i \overline{w_i} \quad (1.12)$$

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

as well as

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

$$= \sum_{i=1}^n v_i \overline{\lambda} \overline{w_i} \quad (1.15)$$

$$= \overline{\lambda} \sum_{i=1}^n v_i \overline{w_i} \quad (1.16)$$

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

Now, we have:

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= \bar{\mu} \langle \mathbf{v}, \mathbf{w} \rangle \quad (1.29)$$

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

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

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

Theorem 1.4.4: Spectral theorem

A symmetric matrix can be diagonalized with an orthonormal matrix \mathbf{S} .

Proof. _____

□

need more
proof

1.4.2 Degree matrix

Definition 1.4.5: 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 1.4.6

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

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

1.4.3 Adjacency matrix

Definition 1.4.7: 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 1.4.8

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 1.4.9

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 1.4.10: Adjacency list

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

Definition 1.4.11: 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 1.4.12

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.

1.4.4 Vertex-edge matrix

Definition 1.4.13: 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 $\mathbf{E}_{n \times m} = [\mathbf{e}_0, \dots, \mathbf{e}_{m-1}]$ be the 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.)

1.4.5 Laplacian matrix

Definition 1.4.14: 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 1.4.15

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

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

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

The degree matrix \mathbf{D} is diagonal and therefore symmetric.

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

Corollary 1.4.16

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 1.4.17

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).

1.4.6 Eigendecomposition

Definition 1.4.18: 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 1.4.19

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

$$\mathbf{Q} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_k \end{bmatrix} \quad (1.31)$$

$$= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{kk} \end{bmatrix} \quad (1.32)$$

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

$$\mathbf{A}\mathbf{Q} = \mathbf{A} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_k \end{bmatrix} \quad (1.33)$$

$$= \begin{bmatrix} \mathbf{A}\mathbf{X}_1 & \mathbf{A}\mathbf{X}_2 & \cdots & \mathbf{A}\mathbf{X}_k \end{bmatrix} \quad (1.34)$$

$$= \begin{bmatrix} \lambda_1 \mathbf{X}_1 & \lambda_2 \mathbf{X}_2 & \cdots & \lambda_k \mathbf{X}_k \end{bmatrix} \quad (1.35)$$

$$= \begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{21} & \cdots & \lambda_k x_{k1} \\ \lambda_1 x_{12} & \lambda_2 x_{22} & \cdots & \lambda_k x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 x_{1k} & \lambda_2 x_{2k} & \cdots & \lambda_k x_{kk} \end{bmatrix} \quad (1.36)$$

$$= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{kk} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix} \quad (1.37)$$

$$= \mathbf{Q}\Lambda \quad (1.38)$$

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 1.4.20

For general positive integer powers $n > 0$,

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

Corollary 1.4.21

The inverse of \mathbf{A} is

$$\mathbf{A}^{-1} = (\mathbf{Q}\Lambda\mathbf{Q}^{-1})^{-1} \tag{1.39}$$

$$= (\mathbf{Q}^{-1})^{-1}(\Lambda)^{-1}(\mathbf{Q})^{-1} \tag{1.40}$$

$$= \mathbf{Q}\Lambda^{-1}\mathbf{Q}^{-1} \tag{1.41}$$

Corollary 1.4.22

The inverse of the diagonal matrix Λ is trivially given by

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

1.4.7 Matrix Exponential**Definition 1.4.23: Matrix Exponential**

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

$$\exp(\mathbf{A}) = e^{\mathbf{A}} \tag{1.42}$$

$$= \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!} \tag{1.43}$$

$$= \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}\mathbf{A}}{2!} + \frac{\mathbf{A}\mathbf{A}\mathbf{A}}{3!} + \cdots \tag{1.44}$$

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

Corollary 1.4.24

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

$$\exp(\Lambda) = e^\Lambda \tag{1.45}$$

$$= \sum_{n=0}^{\infty} \frac{\Lambda^n}{n!} \tag{1.46}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^n \end{bmatrix} \tag{1.47}$$

$$= \begin{bmatrix} \sum_{n=0}^{\infty} \frac{\lambda_1^n}{n!} & 0 & \cdots & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{\lambda_2^n}{n!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{n=0}^{\infty} \frac{\lambda_k^n}{n!} \end{bmatrix} \tag{1.48}$$

$$= \begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_k} \end{bmatrix} \tag{1.49}$$

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

$$\exp(\mathbf{A}) = e^\mathbf{A} \tag{1.50}$$

$$= \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!} \tag{1.51}$$

$$= \sum_{n=0}^{\infty} \frac{\mathbf{Q}\Lambda^n\mathbf{Q}^{-1}}{n!} \tag{1.52}$$

$$= \mathbf{Q} \left(\sum_{n=0}^{\infty} \frac{\Lambda^n}{n!} \right) \mathbf{Q}^{-1} \tag{1.53}$$

$$= \mathbf{Q} e^\Lambda \mathbf{Q}^{-1} \tag{1.54}$$

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

Corollary 1.4.25

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 1.4.26

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

Corollary 1.4.27

If \mathbf{Q} is invertible, then

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

Corollary 1.4.28

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

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

Corollary 1.4.29

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

Chapter 2

Classical Spectral Clustering Algorithm

2.1 Classical Spectral Clustering Algorithm

Bibliography

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