ORIE 4741: Learning with Big Messy Data

Spectral Graph Theory

Mika Sumida
Operations Research and Information Engineering
Cornell

September 15, 2017
Outline

Graph Theory

Spectral Graph Theory

Laplacian regularizer

Spectral Embedding
What is a graph?

A **graph**

- is a collection of nodes that are connected by a set of lines or arrows
- models systems where objects have some pairwise relationship with each other
What is a graph?

A graph

- is a collection of nodes that are connected by a set of lines or arrows
- models systems where objects have some pairwise relationship with each other

Q: What are some examples of graphs in real life?
Examples of graphs: Social network

- Nodes are users
- Edges could be
  - Facebook friendships, LinkedIn connections (undirected)
  - Instagram and Twitter follows (directed)
Examples of graphs: Transportation network

- Subway systems, freight networks
- Roads, bridges, and highway systems
Examples of graphs: Collaboration graphs

- Hollywood graph
- Academic collaborations
Formal definition of graphs

A graph, \( G = (V, E) \), is made up of a

- **Vertex set** \( V = \{v_1, ..., v_n\} \) and an
- **Edge set** \( E = \{e_{ij}\} \)

We say an edge \( e_{ij} \) **connects** vertices \( v_i \) and \( v_j \).
Formal definition of graphs

A graph, \( G = (V, E) \), is made up of a

- **Vertex set** \( V = \{v_1, ..., v_n\} \) and an
- **Edge set** \( E = \{e_{ij}\} \)

We say an edge \( e_{ij} \) **connects** vertices \( v_i \) and \( v_j \).

Two basic types of graphs:

- **Undirected** graphs: edges are sets \( e_{ij} = \{v_i, v_j\} \)
- **Directed** graphs: edges are ordered \( e_{ij} = (v_i, v_j) \)
Formal definition of graphs

A graph, $G = (V, E)$, is made up of a

- **Vertex set** $V = \{v_1, \ldots, v_n\}$ and an
- **Edge set** $E = \{e_{ij}\}$

We say an edge $e_{ij}$ connects vertices $v_i$ and $v_j$.

Two basic types of graphs:

- **Undirected** graphs: edges are sets $e_{ij} = \{v_i, v_j\}$
- **Directed** graphs: edges are ordered $e_{ij} = (v_i, v_j)$

The **degree** of a vertex, $d(v) = \#$ of edges incident to $v$
A graph, $G = (V, E)$, is made up of a

- **Vertex set** $V = \{v_1, ..., v_n\}$ and an
- **Edge set** $E = \{e_{ij}\}$

We say an edge $e_{ij}$ **connects** vertices $v_i$ and $v_j$.

Two basic types of graphs:

- **Undirected** graphs: edges are sets $e_{ij} = \{v_i, v_j\}$
- **Directed** graphs: edges are ordered $e_{ij} = (v_i, v_j)$

The **degree** of a vertex, $d(v) = \#$ of edges incident to $v$

$G$ is **connected** if there is a path between every two vertices in the graph.
Common graphs

- Path graph
Common graphs

- Path graph
- Complete graph
Common graphs

- Path graph
- Complete graph
- Star graph
Common graphs

- Path graph
- Complete graph
- Star graph
- Cycle
After modeling our system as a graph, we can ask about

- maximum degree: finding influential people and celebrities
- finding complete subgraphs (cliques): detecting communities
- minimum cut: sever a communication network into pieces
- shortest paths: routing cars in transportation network
Outline

Graph Theory

Spectral Graph Theory

Laplacian regularizer

Spectral Embedding
Spectral Graph Theory

Spectral graph theory studies properties of graphs through the eigenvalues (spectra) and eigenvectors of associated graph matrices

- Eigenvalues of the Laplacian matrix characterize the connectivity of a graph
- Approximation algorithms for Max Cut
- Spectral clustering
Adjacency Matrix

Encode connections in a graph in a matrix like a spreadsheet.

**Adjacency matrix** is a $|V| \times |V|$ matrix with entries:

$$A(i, j) = \begin{cases} 
1 & \text{if there is an edge } e_{ij} \\
0 & \text{otherwise}
\end{cases}$$
Adjacency Matrix

Encode connections in a graph in a matrix like a spreadsheet

**Adjacency matrix** is a $|V| \times |V|$ matrix with entries:

$$A(i, j) = \begin{cases} 
1 & \text{if there is an edge } e_{ij} \\
0 & \text{otherwise}
\end{cases}$$

Example:
Degree Matrix

**Degree matrix** is a $|V| \times |V|$ matrix with entries:

$$D(i, j) = \begin{cases} 
  d(i) & \text{if } i = j \\
  0 & \text{otherwise}
\end{cases}$$
**Degree Matrix**

**Degree matrix** is a $|V| \times |V|$ matrix with entries:

$$D(i, j) = \begin{cases} d(i) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Example:

\[
\begin{bmatrix}
3 & 0 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 & 0 \\
0 & 0 & 3 & 0 & 0 \\
0 & 0 & 0 & 3 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]
Laplacian Matrix

**Laplacian matrix**, \( L = D - A \)

\[
L(i, j) = \begin{cases} 
  d(i) & \text{if } i = j \\
  -1 & \text{if there is an edge } e_{ij} \\
  0 & \text{otherwise}
\end{cases}
\]

Example:

\[
\begin{pmatrix}
3 & -1 & -1 & -1 & 0 \\
-1 & 2 & -1 & 0 & 0 \\
-1 & -1 & 3 & -1 & 0 \\
-1 & 0 & -1 & 3 & -1 \\
0 & 0 & 0 & -1 & 1
\end{pmatrix}
\]
Laplacians of common graphs

- Path graph

\[ L_{P_4} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \]

- Complete graph

\[ L_{K_4} = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix} \]
Laplacians of common graphs

- **Path graph**

\[
L_{P_4} = \begin{bmatrix}
1 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 1 \\
\end{bmatrix}
\]

- **Complete graph**
Laplacians of common graphs

- Path graph

\[ L_{P_4} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \]

- Complete graph

\[ L_{K_4} = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix} \]
Laplacians of common graphs

- Star graph
Laplacians of common graphs

- Star graph
  \[ L_{S_4} = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \]

- Cycle
Laplacians of common graphs

- **Star graph**

\[ L_{S_4} = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \]

- **Cycle**

\[ L_{C_4} = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix} \]
Matrices as operators

Can think of matrices as functions operating on vectors

\[ M : \mathbb{R}^n \rightarrow \mathbb{R}^n \text{ where } Mv = \begin{bmatrix} -M_1 v \\ -M_2 v \\ \vdots \\ -M_n v \end{bmatrix} \]

What happens if we apply the Laplacian to a vector \( v \)?
Think of $v$ as a distribution of weights or values on the vertices of $G$.

$$(Lv)(i) = L^T_i \, v$$

$$= \sum_{j=1}^{n} L(i, j) v(j)$$

$$= d(i) v(i) + \sum_{j=1, j \neq i}^{n} L(i, j) v(j)$$

$$= d(i) v(i) - \sum_{j: \{i, j\} \in E} v(j)$$

$$= \sum_{j: \{i, j\} \in E} \left( v(i) - v(j) \right)$$
Quadratic Form of the Laplacian

The quadratic form associated with a matrix $M$ is the function

$$f : \mathbb{R}^n \to \mathbb{R} \text{ where } f(\nu) = \nu^T M \nu$$
Quadratic Form of the Laplacian

The **quadratic form** associated with a matrix $M$ is the function

$$f : \mathbb{R}^n \to \mathbb{R} \text{ where } f(v) = v^T M v$$

For the Laplacian matrix,

$$v^T L v = \sum_{i=1}^{n} v(i) \cdot (Lv)(i)$$

$$= \sum_{i=1}^{n} v(i) \sum_{j: \{i,j\} \in E} \left( v(i) - v(j) \right)$$

$$= \sum_{\{i,j\} \in E} v(i) \left( v(i) - v(j) \right) + v(j) \left( v(j) - v(i) \right)$$

$$= \sum_{\{i,j\} \in E} v(i)^2 - 2v(i)v(j) + v(j)^2$$

$$= \sum_{\{i,j\} \in E} \left( v(i) - v(j) \right)^2$$
Laplacian operator

\[(Lv)(i) = \sum_{\{i,j\} \in E} \left( v(i) - v(j) \right) \]

and

\[v^T Lv = \sum_{\{i,j\} \in E} \left( v(i) - v(j) \right)^2\]

- Laplacian operators measure the smoothness of \(v\) across the edges of \(G\)
- If \(v = c \cdot 1\), \(Lv = 0\)
  \(\implies 0\) is an eigenvalue of \(L\) with associated eigenvector \(1\)
- \(v^T Lv \geq 0\) so \(L\) is positive semi-definite
Properties of the Laplacian matrix

- $L$, $D$, and $A$ are all real symmetric matrices
- **Spectral Theorem:** An $n \times n$ real symmetric matrix has $n$ real eigenvalues with $n$ real eigenvectors that form an orthonormal basis
Properties of the Laplacian matrix

- $L$, $D$, and $A$ are all real symmetric matrices
- **Spectral Theorem:** An $n \times n$ real symmetric matrix has $n$ real eigenvalues with $n$ real eigenvectors that form an orthonormal basis
- $L$ is positive semi-definite
  $\Rightarrow$ All eigenvalues of $L$ are non-negative
Properties of the Laplacian matrix

- \( L, D, \) and \( A \) are all real symmetric matrices
- **Spectral Theorem:** An \( n \times n \) real symmetric matrix has \( n \) real eigenvalues with \( n \) real eigenvectors that form an orthonormal basis
- \( L \) is positive semi-definite
  \[ \implies \text{All eigenvalues of } L \text{ are non-negative} \]
- \( L \) has eigenvalues
  \[
  0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n
  \]
Outline

Graph Theory

Spectral Graph Theory

Laplacian regularizer

Spectral Embedding
Smooth regularizer

\[ r(w) = \sum_{i=1}^{d-1} (w_{i+1} - w_i)^2 = \|Sw\|^2 = w^T S^T Sw \]

where \( S \in \mathbb{R}^{(d-1)\times d} \) is the first order difference operator

\[ S(i, j) = \begin{cases} 
1 & j = i \\
-1 & j = i + 1 \\
0 & \text{otherwise}
\end{cases} \]
Smoothed least squares problem

$$\text{minimize } \sum_{i=1}^{n} (y_i - w^T x_i)^2 + \lambda \|Sw\|^2$$

Why smooth?

- can couple coefficients of adjacent features
- allow model to change over space or time
  - example: different years in tax data
**Smoothed least squares problem**

\[
\text{minimize} \quad \sum_{i=1}^{n} (y_i - w^T x_i)^2 + \lambda \|Sw\|^2
\]

Why smooth?

- can couple coefficients of adjacent features
- allow model to change over space or time
  - example: different years in tax data

Can couple any pair of model coefficients, not just \((i, i + 1)\)!
A closer look at the first order difference operator

\[ S^T S = \begin{bmatrix}
1 & -1 & 0 & \cdots & 0 & 0 & 0 \\
-1 & 2 & -1 & \cdots & 0 & 0 & 0 \\
0 & -1 & 2 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 2 & -1 & 0 \\
0 & 0 & 0 & \cdots & -1 & 2 & -1 \\
0 & 0 & 0 & \cdots & 0 & -1 & 1 \\
\end{bmatrix} \]
A closer look at the first order difference operator

\[ S^T S = \begin{bmatrix}
1 & -1 & 0 & \cdots & 0 & 0 & 0 \\
-1 & 2 & -1 & \cdots & 0 & 0 & 0 \\
0 & -1 & 2 & \cdots & 0 & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & 2 & -1 & 0 \\
0 & 0 & 0 & \cdots & -1 & 2 & -1 \\
0 & 0 & 0 & \cdots & 0 & -1 & 1 \\
\end{bmatrix} = L_G(\text{Path}) \]

Laplacian of the path graph

\[ \| Sw \|^2 = w^T L_G(\text{Path}) w \]
Laplacian regularizer

Suppose we have a graph $G$ that encodes relationships between features

- **Product recommendation:** Features are whether a customer bought a certain product. Graph has edges between similar products.

- **Geographic features:** Features are states of residence. Graph has an edge if two states share a border.

- **Time series:** Features are based on years. Graph has edges between consecutive years $(y, y + 1)$.
  - Add in weaker time dependencies by having edges $(y, y + 2)$ with smaller weights.

Laplacian regularizer smooths coefficients over edges of the graph $G$. 
Laplacian Regularized Least Squares

Laplacian regularizer smooths coefficients over edges of the graph $G$.

\[
\text{minimize} \quad \sum_{i=1}^{n} (y_i - w^T x_i)^2 + w^T L_G w
\]

\[
= \sum_{i=1}^{n} (y_i - w^T x_i)^2 + \sum_{\{i,j\} \in E} \left(w(i) - w(j)\right)^2
\]
Outline

Graph Theory

Spectral Graph Theory

Laplacian regularizer

Spectral Embedding
Drawing graphs with eigenvectors

- Laplacian quadratic form measures difference between values of a vector across edges
- For eigenvector of $L$,

$$v^T Lv = \lambda$$

- Eigenvector of $L$ with small eigenvalue has similar $v$-values on adjacent vertices
- Laplacian eigenvectors of low eigenvalue can be used to embed graph into 2-D or 3-D
Spectral Embedding Demo

https://github.com/ORIE4741/demos/spectralGraphTheory.ipynb
References

- Daniel Spielman’s Spectral Graph Theory notes: http://www.cs.yale.edu/homes/spielman/561/
- David Williamson ORIE 6334 notes: https://people.orie.cornell.edu/dpw/orie6334/index.html