Graphs in Machine Learning

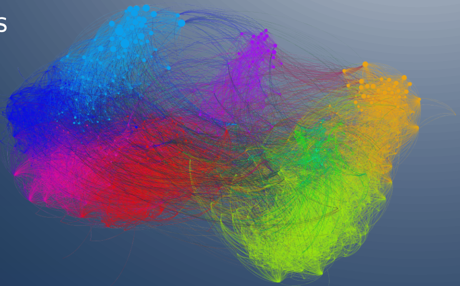
Graph Laplacian Basics

Definition, Properties, Smoothness

Michal Valko

Inria & ENS Paris-Saclay, MVA

Partially based on material by: Ulrike von Luxburg, Gary Miller, Doyle & Schnell, Daniel Spielman



Graph Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ - with a set of **nodes** $\mathcal{V}$ and a set of **edges** $\mathcal{E}$

Graph Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ - with a set of **nodes** $\mathcal{V}$ and a set of **edges** $\mathcal{E}$

$\mathbf{A}$ adjacency matrix

Graph Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ - with a set of **nodes** $\mathcal{V}$ and a set of **edges** $\mathcal{E}$

$\mathbf{A}$	adjacency matrix
$\mathbf{W}$	weight matrix

Graph Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ - with a set of **nodes** $\mathcal{V}$ and a set of **edges** $\mathcal{E}$

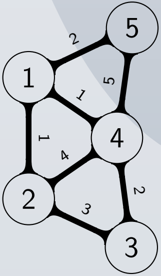
$\mathbf{A}$	adjacency matrix
$\mathbf{W}$	weight matrix
$\mathbf{D}$	(diagonal) degree matrix

Graph Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ - with a set of **nodes** $\mathcal{V}$ and a set of **edges** $\mathcal{E}$

$\mathbf{A}$	adjacency matrix
$\mathbf{W}$	weight matrix
$\mathbf{D}$	(diagonal) degree matrix
$\mathbf{L} = \mathbf{D} - \mathbf{W}$	graph **Laplacian** matrix

$$\mathbf{L} = \begin{pmatrix} 4 & -1 & 0 & -1 & -2 \\ -1 & 8 & -3 & -4 & 0 \\ 0 & -3 & 5 & -2 & 0 \\ -1 & -4 & -2 & 12 & -5 \\ -2 & 0 & 0 & -5 & 7 \end{pmatrix}$$



$\mathbf{L}$ is SDD!

demo: https://dominikschmidt.xyz/spectral-clustering-exp/

Properties of Graph Laplacian

Graph function: a vector $\mathbf{f} \in \mathbb{R}^N$ assigning values to nodes:

$$\mathbf{f} : \mathcal{V}(\mathcal{G}) \to \mathbb{R}.$$

Properties of Graph Laplacian

Graph function: a vector $\mathbf{f} \in \mathbb{R}^N$ assigning values to nodes:

$$\mathbf{f} : \mathcal{V}(\mathcal{G}) \to \mathbb{R}.$$

$$\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$$

Properties of Graph Laplacian

Graph function: a vector $\mathbf{f} \in \mathbb{R}^N$ assigning values to nodes:

$$\mathbf{f} : \mathcal{V}(\mathcal{G}) \to \mathbb{R}.$$

$$\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j \le N} w_{i,j}(f_i - f_j)^2 = S_G(\mathbf{f})$$

Properties of Graph Laplacian

Graph function: a vector $\mathbf{f} \in \mathbb{R}^N$ assigning values to nodes:

$$\mathbf{f} : \mathcal{V}(\mathcal{G}) \to \mathbb{R}.$$

$$\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \frac{1}{2} \sum_{i,j \le N} w_{i,j}(f_i - f_j)^2 = S_G(\mathbf{f})$$

Proof:

$$\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \sum_{i,j \le N} \mathbf{L}_{i,j} f_i f_j = \sum_{i,j \le N} \mathbf{D}_{i,j} f_i f_j - \sum_{i,j \le N} \mathbf{W}_{i,j} f_i f_j = \sum_{i=1}^{N} d_i f_i^2 - \sum_{i,j \le N} w_{i,j} f_i f_j$$

$$= \frac{1}{2}\left(\sum_{i=1}^{N} d_i f_i^2 - 2 \sum_{i,j \le N} w_{i,j} f_i f_j + \sum_{j=1}^{N} d_i f_j^2 \right) = \frac{1}{2} \sum_{i,j \le N} w_{i,j}(f_i - f_j)^2$$

Review: Eigenvalues and Eigenvectors

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

Review: Eigenvalues and Eigenvectors

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

If $(\lambda_1, \mathbf{v}_1)$ are $(\lambda_2, \mathbf{v}_2)$ **eigenpairs** for symmetric $\mathbf{M}$ with $\lambda_1 \neq \lambda_2$ then $\mathbf{v}_1 \perp \mathbf{v}_2$, i.e., $\mathbf{v}_1^\intercal \mathbf{v}_2 = 0$.

Review: Eigenvalues and Eigenvectors

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

If $(\lambda_1, \mathbf{v}_1)$ are $(\lambda_2, \mathbf{v}_2)$ **eigenpairs** for symmetric $\mathbf{M}$ with $\lambda_1 \neq \lambda_2$ then $\mathbf{v}_1 \perp \mathbf{v}_2$, i.e., $\mathbf{v}_1^\intercal\mathbf{v}_2 = 0$.

Proof: $\lambda_1\mathbf{v}_1^\intercal\mathbf{v}_2 = \mathbf{v}_1^\intercal\mathbf{M}\mathbf{v}_2 = \mathbf{v}_1^\intercal\lambda_2\mathbf{v}_2 = \lambda_2\mathbf{v}_1^\intercal\mathbf{v}_2 \implies \mathbf{v}_1^\intercal\mathbf{v}_2 = 0$

Review: Eigenvalues and Eigenvectors

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

If $(\lambda_1, \mathbf{v}_1)$ are $(\lambda_2, \mathbf{v}_2)$ **eigenpairs** for symmetric $\mathbf{M}$ with $\lambda_1 \neq \lambda_2$ then $\mathbf{v}_1 \perp \mathbf{v}_2$, i.e., $\mathbf{v}_1^\mathsf{T}\mathbf{v}_2 = 0$.

Proof: $\lambda_1\mathbf{v}_1^\mathsf{T}\mathbf{v}_2 = \mathbf{v}_1^\mathsf{T}\mathbf{M}\mathbf{v}_2 = \mathbf{v}_1^\mathsf{T}\lambda_2\mathbf{v}_2 = \lambda_2\mathbf{v}_1^\mathsf{T}\mathbf{v}_2 \implies \mathbf{v}_1^\mathsf{T}\mathbf{v}_2 = 0$

If $(\lambda, \mathbf{v}_1)$, $(\lambda, \mathbf{v}_2)$ are eigenpairs for $\mathbf{M}$ then $(\lambda, \mathbf{v}_1 + \mathbf{v}_2)$ is as well.

Review: Eigenvalues and Eigenvectors

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

If $(\lambda_1, \mathbf{v}_1)$ are $(\lambda_2, \mathbf{v}_2)$ **eigenpairs** for symmetric $\mathbf{M}$ with $\lambda_1 \neq \lambda_2$ then $\mathbf{v}_1 \perp \mathbf{v}_2$, i.e., $\mathbf{v}_1^\intercal \mathbf{v}_2 = 0$.

Proof: $\lambda_1 \mathbf{v}_1^\intercal \mathbf{v}_2 = \mathbf{v}_1^\intercal \mathbf{M} \mathbf{v}_2 = \mathbf{v}_1^\intercal \lambda_2 \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^\intercal \mathbf{v}_2 \implies \mathbf{v}_1^\intercal \mathbf{v}_2 = 0$

If $(\lambda, \mathbf{v}_1)$, $(\lambda, \mathbf{v}_2)$ are eigenpairs for $\mathbf{M}$ then $(\lambda, \mathbf{v}_1 + \mathbf{v}_2)$ is as well.

For symmetric $\mathbf{M}$, the **multiplicity** of λ is the dimension of the space of eigenvectors corresponding to λ.

$N \times N$ symmetric matrix has N eigenvalues (w/ multiplicities).

Eigenvalues, Eigenvectors, and Eigendecomposition

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

Vectors $\{\mathbf{v}_i\}_i$ form an **orthonormal** basis with $\lambda_1 \leq \lambda_2 \leq \dots \lambda_N$.

$$\forall i \quad \mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

Eigenvalues, Eigenvectors, and Eigendecomposition

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda\mathbf{v}.$$

Vectors $\{\mathbf{v}_i\}_i$ form an **orthonormal** basis with $\lambda_1 \leq \lambda_2 \leq \ldots \lambda_N$.

$$\forall i \quad \mathbf{M}\mathbf{v}_i = \lambda_i\mathbf{v}_i \qquad \equiv \qquad \mathbf{M}\mathbf{Q} = \mathbf{Q}\boldsymbol{\Lambda}$$

$\mathbf{Q}$ has eigenvectors in columns and $\boldsymbol{\Lambda}$ has eigenvalues on its diagonal.

Eigenvalues, Eigenvectors, and Eigendecomposition

A vector $\mathbf{v}$ is an **eigenvector** of matrix $\mathbf{M}$ of **eigenvalue** λ

$$\mathbf{M}\mathbf{v} = \lambda \mathbf{v}.$$

Vectors $\{\mathbf{v}_i\}_i$ form an **orthonormal** basis with $\lambda_1 \leq \lambda_2 \leq \dots \lambda_N$.

$$\forall i \quad \mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{v}_i \qquad \equiv \qquad \mathbf{M}\mathbf{Q} = \mathbf{Q}\boldsymbol{\Lambda}$$

$\mathbf{Q}$ has eigenvectors in columns and $\boldsymbol{\Lambda}$ has eigenvalues on its diagonal.

Right-multiplying $\mathbf{M}\mathbf{Q} = \mathbf{Q}\boldsymbol{\Lambda}$ by $\mathbf{Q}^\intercal$ we get the **eigendecomposition** of $\mathbf{M}$:

$$\mathbf{M} = \mathbf{M}\mathbf{Q}\mathbf{Q}^\intercal = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\intercal = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^\intercal$$

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{L}$ positive semi-definite

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{L}$ positive semi-definite $\mathbf{f}^\intercal \mathbf{L} \mathbf{f} = \frac{1}{2} \sum_{i,j \leq N} w_{i,j} (f_i - f_j)^2$

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{L}$ positive semi-definite $\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j\leq N} w_{i,j}(f_i - f_j)^2$

Recall: If $\mathbf{L}\mathbf{f} = \lambda\mathbf{f}$ then λ is an **eigenvalue** (of the Laplacian).

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{L}$ positive semi-definite $\mathbf{f}^{\mathsf{T}}\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$

Recall: If $\mathbf{L}\mathbf{f} = \lambda\mathbf{f}$ then λ is an **eigenvalue** (of the Laplacian).

The smallest eigenvalue of $\mathbf{L}$ is 0. Corresponding eigenvector: $\mathbf{1}_N$.

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{L}$ positive semi-definite $\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j\leq N} w_{i,j}(f_i - f_j)^2$

Recall: If $\mathbf{L}\mathbf{f} = \lambda\mathbf{f}$ then λ is an **eigenvalue** (of the Laplacian).

The smallest eigenvalue of $\mathbf{L}$ is 0. Corresponding eigenvector: $\mathbf{1}_N$.

All eigenvalues are non-negative reals $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$.

$\mathbf{M} = \mathbf{L}$: Properties of Graph Laplacian

We can assume **non-negative weights**: $w_{ij} \geq 0$.

$\mathbf{L}$ is symmetric

$\mathbf{L}$ positive semi-definite $\mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j\leq N} w_{i,j}(f_i - f_j)^2$

Recall: If $\mathbf{L}\mathbf{f} = \lambda\mathbf{f}$ then λ is an **eigenvalue** (of the Laplacian).

The smallest eigenvalue of $\mathbf{L}$ is 0. Corresponding eigenvector: $\mathbf{1}_N$.

All eigenvalues are non-negative reals $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$.

Self-edges do not change the value of $\mathbf{L}$.

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Proof: If $(0, \mathbf{f})$ is an eigenpair then $0 = \frac{1}{2}\sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$.

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Proof: If $(0, \mathbf{f})$ is an eigenpair then $0 = \frac{1}{2}\sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$. Therefore, $\mathbf{f}$ is constant on each connected component.

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Proof: If $(0, \mathbf{f})$ is an eigenpair then $0 = \frac{1}{2}\sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$. Therefore, $\mathbf{f}$ is constant on each connected component. If there are k components, then $\mathbf{L}$ is k-block-diagonal:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & \ddots & \\ & & & \mathbf{L}_k \end{bmatrix}$$

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Proof: If $(0, \mathbf{f})$ is an eigenpair then $0 = \frac{1}{2}\sum_{i,j\leq N} w_{i,j}(f_i - f_j)^2$. Therefore, $\mathbf{f}$ is constant on each connected component. If there are k components, then $\mathbf{L}$ is k-block-diagonal:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & \ddots & \\ & & & \mathbf{L}_k \end{bmatrix}$$

For block-diagonal matrices: the spectrum is the union of the spectra of $\mathbf{L}_i$

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Proof: If $(0, \mathbf{f})$ is an eigenpair then $0 = \frac{1}{2}\sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$. Therefore, $\mathbf{f}$ is constant on each connected component. If there are k components, then $\mathbf{L}$ is k-block-diagonal:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & \ddots & \\ & & & \mathbf{L}_k \end{bmatrix}$$

For block-diagonal matrices: the spectrum is the union of the spectra of $\mathbf{L}_i$ (eigenvectors of $\mathbf{L}_i$ padded with zeros elsewhere).

Properties of Graph Laplacian

The multiplicity of eigenvalue 0 of $\mathbf{L}$ equals to the number of connected components. The eigenspace of 0 is spanned by the components' indicators.

Proof: If $(0, \mathbf{f})$ is an eigenpair then $0 = \frac{1}{2} \sum_{i,j \leq N} w_{i,j}(f_i - f_j)^2$. Therefore, $\mathbf{f}$ is constant on each connected component. If there are k components, then $\mathbf{L}$ is k-block-diagonal:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & \ddots & \\ & & & \mathbf{L}_k \end{bmatrix}$$

For block-diagonal matrices: the spectrum is the union of the spectra of $\mathbf{L}_i$ (eigenvectors of $\mathbf{L}_i$ padded with zeros elsewhere).

For $\mathbf{L}_i$ $(0, \mathbf{1}_{|V_i|})$ is an eigenpair, hence the claim.

Michal Valko

michal.valko@inria.fr

Inria & ENS Paris-Saclay, MVA

https://misovalko.github.io/mva-ml-graphs.html