Graphs in Machine Learning

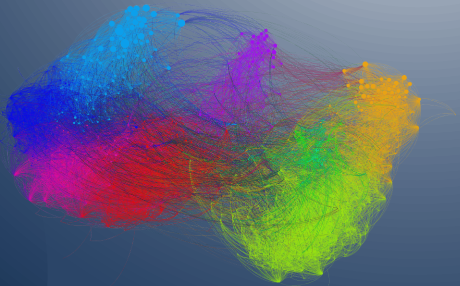
Spectral Clustering: NCut

Normalized Cut Approximation

Michal Valko

Inria & ENS Paris-Saclay, MVA

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



Spectral Clustering: Approximating NCut

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Spectral Clustering: Approximating NCut

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Define graph function $\mathbf{f}$ for cluster membership of NCut:

$$f_i = \begin{cases} \sqrt{\frac{\mathrm{vol}(B)}{\mathrm{vol}(A)}} & \text{if } V_i \in A, \\ -\sqrt{\frac{\mathrm{vol}(A)}{\mathrm{vol}(B)}} & \text{if } V_i \in B. \end{cases}$$

Spectral Clustering: Approximating NCut

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Define graph function $\mathbf{f}$ for cluster membership of NCut:

$$f_i = \begin{cases} \sqrt{\frac{\mathrm{vol}(B)}{\mathrm{vol}(A)}} & \text{if } V_i \in A, \\ -\sqrt{\frac{\mathrm{vol}(A)}{\mathrm{vol}(B)}} & \text{if } V_i \in B. \end{cases}$$

$$(\mathbf{Df})^\mathsf{T} \mathbf{1}_n = 0$$

Spectral Clustering: Approximating NCut

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Define graph function $\mathbf{f}$ for cluster membership of NCut:

$$f_i = \begin{cases} \sqrt{\frac{\mathrm{vol}(B)}{\mathrm{vol}(A)}} & \text{if } V_i \in A, \\ -\sqrt{\frac{\mathrm{vol}(A)}{\mathrm{vol}(B)}} & \text{if } V_i \in B. \end{cases}$$

$$(\mathbf{Df})^\mathsf{T} \mathbf{1}_n = 0 \qquad \mathbf{f}^\mathsf{T} \mathbf{D} \mathbf{f} = \mathrm{vol}(\mathcal{V})$$

Spectral Clustering: Approximating NCut

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Define graph function $\mathbf{f}$ for cluster membership of NCut:

$$f_i = \begin{cases} \sqrt{\frac{\mathrm{vol}(B)}{\mathrm{vol}(A)}} & \text{if } V_i \in A, \\ -\sqrt{\frac{\mathrm{vol}(A)}{\mathrm{vol}(B)}} & \text{if } V_i \in B. \end{cases}$$

$$(\mathbf{Df})^{\mathsf{T}} \mathbf{1}_n = 0 \qquad \mathbf{f}^{\mathsf{T}} \mathbf{Df} = \mathrm{vol}(\mathcal{V}) \qquad \mathbf{f}^{\mathsf{T}} \mathbf{Lf} = \mathrm{vol}(\mathcal{V}) \mathrm{NCut}(A, B)$$

Spectral Clustering: Approximating NCut

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Define graph function $\mathbf{f}$ for cluster membership of NCut:

$$f_i = \begin{cases} \sqrt{\frac{\mathrm{vol}(B)}{\mathrm{vol}(A)}} & \text{if } V_i \in A, \\ -\sqrt{\frac{\mathrm{vol}(A)}{\mathrm{vol}(B)}} & \text{if } V_i \in B. \end{cases}$$

$$(\mathbf{Df})^\mathsf{T}\mathbf{1}_n = 0 \qquad \mathbf{f}^\mathsf{T}\mathbf{Df} = \mathrm{vol}(\mathcal{V}) \qquad \mathbf{f}^\mathsf{T}\mathbf{Lf} = \mathrm{vol}(\mathcal{V})\mathrm{NCut}(A, B)$$

objective function of spectral clustering (NCut)

$$\min_{\mathbf{f}} \mathbf{f}^\mathsf{T}\mathbf{Lf} \quad \text{s.t.} \quad f_i \in \mathbb{R}, \quad \mathbf{Df} \perp \mathbf{1}_N, \quad \mathbf{f}^\mathsf{T}\mathbf{Df} = \mathrm{vol}(\mathcal{V})$$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{f}} \mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} \quad \text{s.t.} \quad f_i \in \mathbb{R}, \quad \mathbf{D}\mathbf{f} \perp \mathbf{1}_N, \quad \mathbf{f}^\mathsf{T}\mathbf{D}\mathbf{f} = \mathrm{vol}(\mathcal{V})$$

Can we apply Rayleigh-Ritz now?

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{f}} \mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} \quad \text{s.t.} \quad f_i \in \mathbb{R}, \quad \mathbf{D}\mathbf{f} \perp \mathbf{1}_N, \quad \mathbf{f}^\mathsf{T}\mathbf{D}\mathbf{f} = \mathrm{vol}(\mathcal{V})$$

Can we apply Rayleigh-Ritz now? Define $\mathbf{w} = \mathbf{D}^{1/2}\mathbf{f}$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{f}} \mathbf{f}^{\mathsf{T}}\mathbf{L}\mathbf{f} \quad \text{s.t.} \quad f_i \in \mathbb{R}, \quad \mathbf{D}\mathbf{f} \perp \mathbf{1}_N, \quad \mathbf{f}^{\mathsf{T}}\mathbf{D}\mathbf{f} = \mathrm{vol}(\mathcal{V})$$

Can we apply Rayleigh-Ritz now? Define $\mathbf{w} = \mathbf{D}^{1/2}\mathbf{f}$

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^{\mathsf{T}}\mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \mathbf{w} \perp \mathbf{D}^{1/2}\mathbf{1}_N, \|\mathbf{w}\|^2 = \mathrm{vol}(\mathcal{V})$$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{f}} \mathbf{f}^\mathsf{T}\mathbf{L}\mathbf{f} \quad \text{s.t.} \quad f_i \in \mathbb{R}, \quad \mathbf{D}\mathbf{f} \perp \mathbf{1}_N, \quad \mathbf{f}^\mathsf{T}\mathbf{D}\mathbf{f} = \mathrm{vol}(\mathcal{V})$$

Can we apply Rayleigh-Ritz now? Define $\mathbf{w} = \mathbf{D}^{1/2}\mathbf{f}$

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \mathbf{w} \perp \mathbf{D}^{1/2}\mathbf{1}_N, \|\mathbf{w}\|^2 = \mathrm{vol}(\mathcal{V})$$

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{L}_{\mathrm{sym}}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\mathrm{sym}}}, \quad \|\mathbf{w}\|^2 = \mathrm{vol}(\mathcal{V})$$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{L}_{\mathrm{sym}}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\mathrm{sym}}}, \quad \|\mathbf{w}\| = \mathrm{vol}(\mathcal{V})$$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T} \mathbf{L}_{\mathrm{sym}} \mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\mathrm{sym}}}, \quad \|\mathbf{w}\| = \mathrm{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz?

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T} \mathbf{L}_{\mathrm{sym}} \mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\mathrm{sym}}}, \quad \|\mathbf{w}\| = \mathrm{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz? $\mathbf{w} = \mathbf{v}_{2,\mathbf{L}_{\mathrm{sym}}}$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{L}_{\text{sym}}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\text{sym}}}, \quad \|\mathbf{w}\| = \text{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz? $\mathbf{w} = \mathbf{v}_{2,\mathbf{L}_{\text{sym}}}$ $\mathbf{f} =$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T} \mathbf{L}_{\text{sym}} \mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\text{sym}}}, \quad \|\mathbf{w}\| = \text{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz? $\mathbf{w} = \mathbf{v}_{2,\mathbf{L}_{\text{sym}}}$ $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{w}$

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{L}_{\text{sym}}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\text{sym}}}, \quad \|\mathbf{w}\| = \text{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz? $\mathbf{w} = \mathbf{v}_{2,\mathbf{L}_{\text{sym}}}$ $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{w}$

$\mathbf{f}$ is a the second eigenvector of $\mathbf{L}_{\text{rw}}$!

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{L}_{\text{sym}}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\text{sym}}}, \quad \|\mathbf{w}\| = \text{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz? $\mathbf{w} = \mathbf{v}_{2,\mathbf{L}_{\text{sym}}}$ $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{w}$

$\mathbf{f}$ is a the second eigenvector of $\mathbf{L}_{\text{rw}}$!

tl;dr: Get the second eigenvector of $\mathbf{L}/\mathbf{L}_{\text{rw}}$ for RatioCut/NCut.

Spectral Clustering: Approximating NCut

objective function of spectral clustering (NCut)

$$\min_{\mathbf{w}} \mathbf{w}^\mathsf{T}\mathbf{L}_{\text{sym}}\mathbf{w} \quad \text{s.t.} \quad w_i \in \mathbb{R}, \quad \mathbf{w} \perp \mathbf{v}_{1,\mathbf{L}_{\text{sym}}}, \quad \|\mathbf{w}\| = \text{vol}(\mathcal{V})$$

Solution by Rayleigh-Ritz? $\mathbf{w} = \mathbf{v}_{2,\mathbf{L}_{\text{sym}}}$ $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{w}$

$\mathbf{f}$ is a the second eigenvector of $\mathbf{L}_{\text{rw}}$!

tl;dr: Get the second eigenvector of $\mathbf{L}/\mathbf{L}_{\text{rw}}$ for RatioCut/NCut.

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

Spectral Clustering: Approximation

These are all approximations. How bad can they be?

Spectral Clustering: Approximation

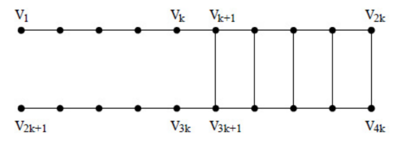
These are all approximations. How bad can they be?

Pretty bad.

Spectral Clustering: Approximation

These are all approximations. How bad can they be?

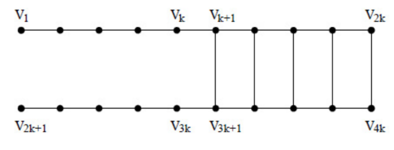
Pretty bad. Example: cockroach graphs



Spectral Clustering: Approximation

These are all approximations. How bad can they be?

Pretty bad. Example: cockroach graphs



No efficient approximation exist. Other relaxations possible.

https://www.cs.cmu.edu/~glmiller/Publications/Papers/GuMi95.pdf

Michal Valko

michal.valko@inria.fr

Inria & ENS Paris-Saclay, MVA

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