Spectral Clustering of Signed Graphs via Matrix Power Means

Pedro Mercado, Francesco Tudisco and Matthias Hein

UNIVERSITÄT DES SAARLANDES

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

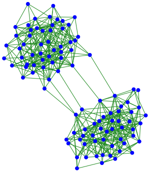
ICML 2019, Long Beach, USA

Poster #190

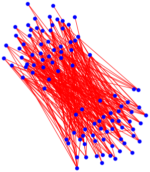
Our Goal: Extend Spectral Clustering to Graphs With Both Positive and Negative Edges

- **Positive Edges:** encode friendship, similarity, proximity, trust
- **Negative Edges:** encode enmity, dissimilarity, conflict, distrust

$G^{\pm} = ($



$,$



$)$

A signed graph is the pair $G^{\pm} = (G^{+}, G^{-})$ where

$G^{+} = (V, W^{+})$ encodes **positive** relations, and

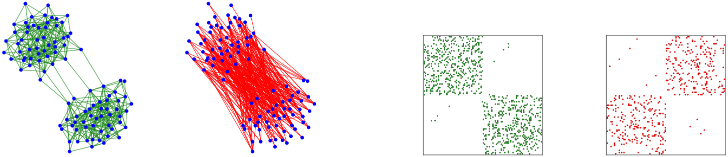
$G^{-} = (V, W^{-})$ encodes **negative** relations

Clustering of Signed Graphs

Given: an undirected signed graph $G^{\pm} = (G^{+}, G^{-})$

Goal : partition the graph such that

- edges **within** the same group have **positive** weights
- edges **between** different groups have **negative** weights



G^{+} G^{-} W^{+} W^{-}

Our Goal: define an operator that <u>blends</u> the information of (G^{+}, G^{-}) such that the smallest eigenvectors are informative.

Our Goal: define an operator that blends the information of (G^+, G^-) such that the smallest eigenvectors are informative.

State of the art approaches:

$$\mathbf{L}_{\mathbf{SR}} = \mathbf{L}^+ + \mathbf{Q}^- \qquad \text{(Kunegis, 2010)}$$

$$\mathbf{L}_{\mathbf{BR}} = \mathbf{L}^+ + \mathbf{W}^- \qquad \text{(Chiang, 2012)}$$

$$\mathbf{H} = (\alpha - 1)\mathbf{I} - \sqrt{\alpha}(\mathbf{W}^+ - \mathbf{W}^-) + \mathbf{D}^+ + \mathbf{D}^- \qquad \text{(Saade, 2015)}$$

Current methods are arithmetic means of Laplacians

The **power mean** of non-negative scalars a, b, and $p \in \mathbb{R}$:

$$m_p(a, b) = \left(\frac{a^p + b^p}{2} \right)^{1/p}$$

Particular cases of the scalar power mean are:

$p \to -\infty$	$p = -1$	$p \to 0$	$p = 1$	$p \to \infty$
$\min\{a, b\}$	$2\left(\frac{1}{a} + \frac{1}{b}\right)^{-1}$	$\sqrt{ab}$	$(a+b)/2$	$\max\{a, b\}$
minimum	harmonic mean	geometric mean	arithmetic mean	maximum

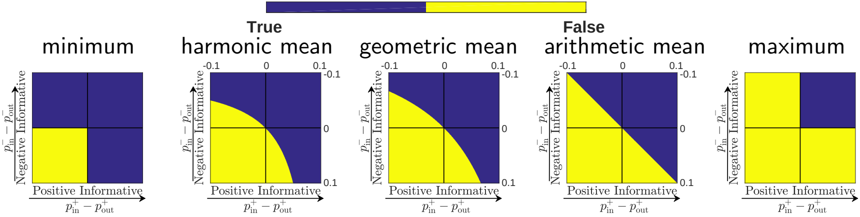
We introduce the **Signed Power Mean Laplacian** as an alternative to **blend the information** of the signed graph $G^{\pm}$:

$$\mathbf{L_p} = \left(\frac{\left(\mathbf{L}^{+}_{\mathbf{sym}}\right)^p + \left(\mathbf{Q}^{-}_{\mathbf{sym}}\right)^p}{2} \right)^{1/p}$$

Analysis in the Stochastic Block Model

Theorem (loosely stated): The Signed Power Mean Laplacian L_p with $p \leq 0$ is better than arithmetic mean approaches in expectation.

Recovery of Clusters in Expectation



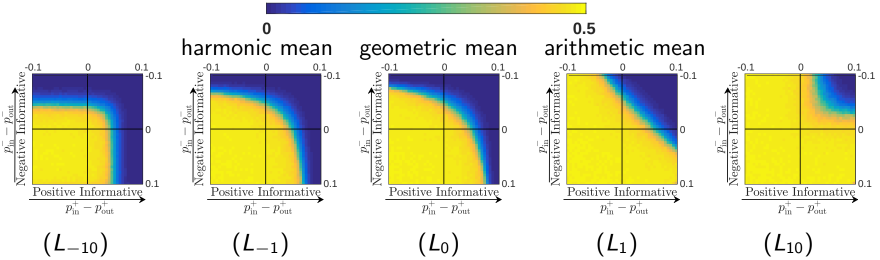
$(\mathcal{L}_{-\infty})$ $(\mathcal{L}_{-1})$ $(\mathcal{L}_{0})$ $(\mathcal{L}_{1})$ $(\mathcal{L}_{\infty})$

Spectral Clustering of Signed Graphs   Poster #190

Analysis in the Stochastic Block Model

Theorem (loosely stated): The Signed Power Mean Laplacian L_p with $p \leq 0$ is better than arithmetic mean approaches in expectation.

Average Clustering Error



Theorem (loosely stated): with high probability eigenvalues and eigenvectors of L_p concentrate around those of the expected Signed Power Mean Laplacian $\mathcal{L}_p$