The University of Hong Kong

FYP18026 Final Year Project

Project Plan

Diffusion Process with Mediator on Hypergraph

Author
Gao Yuan 3035234545

Supervisor
Dr. Hubert Chan

September 2018

Contents

1 Introduction

The spectral theory studies the spectral properties of graphs and their relation to the combinatorics properties of graph. Central to this rich theory, the Cheeger's inequality bounds the expansion or conductance of the graph by the second least eigenvalue of the Laplacian of the graph[1]. More specifically, for a graph $G = (V, E)$, its expansion is defined to be:

$$\phi_G := \min_{S \subset V} \frac{|\partial S|}{\min\{vol(S), vol(\bar{S})\}},$$

where $vol(S)$ is the sum of the degree of vertices in S and ∂S is the cut of S. The Cheeger's inequality[1] is as the following:

$$\frac{\lambda_2}{2} \leq \phi_G \leq \sqrt{2\lambda_2},$$

where λ_2 is the second least eigenvalue of the normalized Laplacian $\mathcal{L}_G$. This inequality can be utilized in the design of various algorithms.

Recently, the Cheeger's inequality was generalized to the case of hypergraph with notions of hypergraph expansion and hypergraph Laplacian[2][3]. This hypergraph Laplacian, denoted by L_ω, is defined to be an operator that determines a diffusion process with mediators over hypergraphs. In [2][3], it is shown that this Laplacian operator L_ω possesses a second least eigenvalue λ_2 that satisfy an inequality similar to that of Cheeger's:

$$\frac{\lambda_2}{2} \leq \phi_H \leq \sqrt{2\lambda_2},$$

where ϕ_H is the expansion of hypergraph. This result shows that the spectral properties of hypergraph Laplacian can indeed be linked to combinatorics properties of hypergraph. However, unlike graph Laplacians, which are symmetric linear operators, the hypergraph Laplacians are non-linear and non-smooth, which makes its spectrum hard to characterize. Hypergraph Laplacian is shown to have a trivial eigenvalue and a second least eigenvalue[2][3] and in this project the conjecture that the hypergraph Laplacian possesses a maximum eigenvalue is proved. Furthermore, the maximum eigenvalue of the hypergraph Laplacian is related to the approximation of the hypergraph max cut, the computation of which is NP-hard. It is thus interesting to consider the approximation of the maximum eigenvalue and its corresponding eigenvectors. One of the goals of the project would therefore be studying possible approximation algorithms of the maximum eigenvalue of the hypergraph Laplacian.

A more generalized notion of Laplacian can be defined with respect to submodular transformations[4]. Hypergraph Laplacians and graph Laplacians can be viewed as special cases of the generalized Laplacian by considering different submodular transformations corresponding to hypergraphs or graphs. This more abstract construction gives rise to some novel Cheeger's inequalities under the same unifying framework. It is thus interesting to consider the spectral properties of this Laplacian. In [4], the generalized Laplacian is shown to possess two eigenpairs, similar to the case of hypergraph Laplacians. One possible research direction along this line would therefore be the study of other eigenpairs of the generalized Laplacians. The conjecture is that the generalized Laplacian also possesses a maximum eigenvalue. Hence a proof for this conjecture shall be studied in this project.

One particular obstacle of the application of spectral theory of hypergraphs is the size of the edge set of hypergraphs. For a hypergraph with n vertices, the number of the edges of the hypergraph can be of the order $O(2^n)$, which can cause the evaluation of the hypergraph Laplacian and the induced quadratic form intractable. In [5], a spectral sparsifier of hypergraphs is proposed, with a theoretical guarantee that the number of edges of the sparsified hypergraph is of the order $O(n^3 logn/\epsilon^2)$ with high probability, where ϵ is the error of the quadratic form of the sparsified hypergraph. It might be interesting to see if it is possible to further reduce the number of the edges in the sparsified hypergraph. Application-wise speaking, as this sparsifier makes the computations involving the quadratic form of the Laplacian of dense hypergraphs feasible, experiments could be carried out to test how this sparsifier might improve the performances of algorithms that involves hypergraph Laplacian, for example semi-supervised learning on hypergraphs[6].

The remainder of this project plan is as follows. First, I will give more explanations on the notions of hypergraph Laplacian, the induced quadratic form and other related notions. Then I will define the problems to be studied and consider how the study of them can benefit the applications of the spectral theory of hypergraphs. The project plan will close with a discussion on the methodology and a project schedule.

2 Problem Definitions and Goals

In this project we consider edge-weighted graphs $H = (V, E, \omega)$ where ω is a consistent weight function. We assume that all vertex weights are positive. Given $f \in \mathbb{R}^V$, $f_u, u \in V$ is the coordinate corresponding to $u \in V$. In addition, the space $\mathbb{R}^V$ is endowed with the inner product $\langle \cdot, \cdot \rangle_\omega$ defined as $\langle f, f \rangle_\omega = \langle f, Wf \rangle$ where W is the diagonal matrix of the weights of the vertices. This inner product space is called the density space.

2.1 Generalized Quadratic Form and Rayleigh Quotient

In[3], the notion of diffusion operator with mediator is defined. For each edge $e \in E$, let $[e] = e \cup \{0\}$, where 0 does not correspond to any vertex. A set of constants β^e is defined over $[e]$ with β^e_j corresponding to all $j \in [e]$. Then the following associated generalized quadratic form is considered:

$$Q(f) := \sum_{e\in E} \omega_e \{\beta^e_0 \max_{s,i\in e}(f_s - f_i)^2 + \sum_{j\in e} \beta^e_j [(\max_{s\in e} f_s - f_j)^2 + (\min_{i\in e} f_i - f_j)^2]\}.$$

Also, for $f \neq \mathbf{0} \in \mathbb{R}^V$, $\frac{Q(f)}{\sum_{u\in V} \omega_u f_u^2}$ is defined to be its discrepancy ratio.

As in[3], a diffusion process (with mediator) can be defined according a set of rules. Then by considering the differentiation of the density of the measure over the vertices in the diffusion process, the Laplacian L_ω can be defined by $L_\omega f = \frac{df}{dt}$, where t is the time of the diffusion process. The detailed explanation of the diffusion process with mediator and the proof of its existence can be found in [2] and [3]. It is also shown that the Rayleigh Quotient of the hypergraph Laplacian L_ω, defined as $R(f) = \frac{\langle f, L_\omega f\rangle_\omega}{\langle f, f\rangle_\omega}$ for any $\mathbf{0} \neq f \in \mathbb{R}^V$, coincides with the descrepancy ratio of f.

2.2 Eigenpairs of the Hypergraph Laplacian

It is obvious that the hypergraph Laplacian obtains a trivial eigenpair. More specifically, $\mathbf{1} \in \mathbb{R}^V$ is an eigenvector of the Laplacian with 0 being its corresponding eigenvalue. It is not as trivial to find other eigenpairs of the Laplacian. In [2][3], it is shown that $\lambda_2 := \min_{\mathbf{1}\perp f\in\mathbb{R}^V} R(f)$ is an eigenvalue with any $\mathbf{1}\perp f$ attaining this value being an eigenvector.

It was left as an open question whether there are other eigenpairs for the hypergraph Laplacian or not. It is now known that the maximum of the Rayleigh quotient, i.e. $\lambda^* = \max_{\mathbf{1}\perp f\in\mathbb{R}^V} R(f)$, is the maximum eigenvalue of the hypergraph Laplacian, with any $\mathbf{1}\perp f$ attaining this value being an eigenvector. This eigenpair is closely related to the max-cut of hypergraphs, the computation of which is NP-hard. Hence knowing the maximum eigenvalue and its corresponding eigevector might be useful for the approximation of the max cut, which requires an approximation algorithm for the maximum of the Rayleigh quotient of hypergraphs. Therefore, one of the objectives of this project will be on the approximation of this eigenpair. In [2], a semi-definite programming (SDP) relaxation and a rounding algorithm is proposed for the approximation of the procedural minimizer of the discrepancy ration of hypergraphs. The first two eigenpairs of the hypergraph Laplacian are in fact the first two procedural minimizers of the discrepancy ratio. One possible idea for the approximation of the maximum eigenvalue of the hypergraph Laplacian might be formulating a similar SDP relaxation and rounding algorithm. However, it is not known if a similar SDP relaxation is possible since the two approximation problems are subtly different. This project will investigate into this idea and perhaps explore various techniques on solving SDP.

2.3 Spectral Sparsifier of Hypergraph

The size of the edge size of hypergraphs can be of order 2^n. This might cause the evaluation of the hypergraph Laplacian and Rayleigh Quotient (or the genralized quadratic form) intractable. It is therefore important to consider the spectral sparisifier of hypergraph. A subgraph H' of the hypergraph H is an ϵ-spectral sparsifier of H if:

$$(1-\epsilon)Q_{H'}(f) \le Q_H(f) \le (1+\epsilon)Q_{H'}(f)$$

for every $f \in \mathbb{R}^V$. A randomized algorithm is proposed in [5] that can generate a sparsifier with $O(n^3 logn/\epsilon^2)$ edges with high probability for any hypergraph, which makes the computation of various problems that involves the quadratic form of hypergraphs feasible. Some possible directions along this line of thought might be conducting some experiments to quantify emperically the benefits that the hypergraph sparsifier can bring to different computational models. It might be interesting to study the computational models that arises in natural scenarios where the quadratic form of dense hypergraphs are needed.

One particular example of the possible application of the sparsifier might be the semi-supervised learning on hypergraphs[6], which in principle can be accelerated by spectral sparsifiers[5]. In [6], a learning task on (directed) hypergraph is considered. Labels $x^*(u)$ are given for some vertices $u \in L \subset V$ and the task is to predict the labels of vertices in $V\backslash L$. This problem is then transformed into an optimization problem with the generalized quadratic form as the objective. A subgradient method is proposed to solve this optimization problem[6]. Since clearly this optimization problem involves the generalized quadratic form, applying the spectral sparsifier of hypergraph in this case should be able to speed up the computations significantly, especially when the hypergraph model is dense[5]. Then an experimental work of this project will be on finding dense hypergraph models in the scenario of semi-supervised learning and experiment the spectral sparsifier with it.

Another possible direction might be improving the sparsifier proposed by [5]. Note that the current sparsifier is of order $O(n^3 logn/\epsilon^2)$. However, a hypergraph with edges that contains no more than 3 vertices has no more than $O(n^3)$ edges, which means that the proposed sparsifier will not sparsify this hypergraph at all. Hence it would be interesting to consider sparsifiers that might give better theoretical guarantees.

2.4 Generalized Laplacian and Its Eigenpairs

The notion of Laplacican can be further generalized for submodular transformations, which is the product of submodular functions over the power set of some set V, or equivalently $\{0,1\}^V$[4]. That is to say, a function $F : \{0,1\}^V \to \mathbb{R}^E$ is said to be a submodular transformation if each of its component function is submodular. A set function $F_e : \{0,1\}^V \to \mathbb{R}$ is a submodular function if $F_e(S)+F_e(T) \ge F_e(S\cap T)+F_e(S\cup T), \forall S,T \subset V$. For each submodular F_e, there exists a Lovász extension $f_e : [0,1]^V \to \mathbb{R}$ which is defined to be the convex closure of the submodular function F_e. More specifically, [4] considers the submodular polyhedron $P(F)$ and the base polytope $B(F)$ of F:

$$P(F) = \{x \in \mathbb{R}^V | \textstyle\sum_{v\in S} x(v) \le F(S), \forall S \in V\} \text{ and } B(F) = \{x \in P(F) | \textstyle\sum_{v\in V} x(v) = F(V)\}.$$

With $P(F)$ and $B(F)$, the Lovász extension of $f_e : \mathbb{R}^V \to \mathbb{R}$ of a submodular function $F_e : \{0,1\}^V \to \mathbb{R}$ is:

$$f_e(x) = \max_{\omega\in B(F_e)} \langle \omega, x \rangle.$$

Note that $\partial f_e = argmax_{\omega \in B(F_e)} \langle \omega, x \rangle$.

In [4], the notion of Laplacian operator $L_F : \mathbb{R}^V \to 2^{\mathbb{R}^V}$ is defined for the submodular transformation F as follows:

$$L_F(x) := \{\sum_{e \in E} \langle \omega_e, x \rangle | \omega_e \in \partial f_e(x)\},$$

where $\partial f(x)$ is the subgradient of f_e at x. The Lovász extension of f of a submodular function is defined to be the product of Lovász extensions of each component functions of F. Then [4] shows that the Rayleigh quotient $R_F : \mathbb{R}^V \to \mathbb{R}$ of the generalized Laplacian L_F for submodular transformation F is:

$$R_F(x) := \frac{\langle x, L_F(x) \rangle}{\langle x, x \rangle} = \frac{\|f(x)\|_2^2}{\|x\|_2^2}, \text{ for } x \neq \mathbf{0},$$

where $\|\cdot\|_2$ is the 2-norm over $\mathbb{R}^V$. Graph Laplacian and hypergraph Laplacian are special cases of this generalized Laplacian. Like the case of hypergraph Laplacian, it is known that L_F admitts two eigenpairs, one of them being trivial, and the other being the minimizer of the Rayleigh quotient subject to the condition that $x \neq \mathbf{0}$ and $x \perp \mathbf{1}$[4]. It is not yet known if there exist other eigenpairs of the generalized Laplacian. Therefore, I will attempt to generalize the proof for the existence of the maximum eigenvalue of the hypergraph Laplacian to the case of generalized Laplacian. If such generalization is indeed true, then the study of the approximation algorithm of this possible eigenpair will be in order. This will provide a unifying framework for the previous works such that results concerning the second least and maximum eigenvalues of the graph Laplacian and hypergraph Laplacian can all be derived from that of the generalized Laplacian by considering different submodular transformations.

3 Methodology

As this point, the project is very theoretical and mathematical in nature. A significant amount of effort will be put on proving mathematical conjectures. The rest of the project will mostly be on the study of various approximation algorithms, which will involve theoretical analysis of the performances. Thus this project will not be implementation-oriented. As such, a significant amount of time will be spent on literature reviews, instead of programming and testing. I will try to explore extensively on the mathematical ideas and tools that might be helpful for the project.

4 Project Schedule

Due to the nature of the project, it is not very practical to put down a precise schedule for the project. Coming up with correct proofs for the problems discussed or finding out efficient approximation algorithms might take much more or if lucky, less time than expected. Hence a rough time frame is given as follows:

1. **October**: Study relevant papers on the problems discussed above. Read extensively on mathematics that might be useful, including optimization, combinatorics and algebra and so on. Investigate into possible ideas and solutions to the problems above. Start building up the final project report.

2. **November and December**: Focus more on problem solving. If there are theoretical progresses, validate them. Identify relevant problems that are also of interests and if time allows, work on them. Add any new findings to the project report.

3. **January to March**: If there are any progress in the theoretical work, continue investigating them. If not, some time should be put into experimental work, including the experiments on possible applications of the spectral sparsifier of graphs. The experiments will involve gathering of data (as represented in dense hypergraphs) and implementation of the existing sparsification algorithms. Finish the draft project report.

4. **April and Onward**: Wrap up the project. Identify the problems solved and problems unsolved. Identify possible future works. Finalize the project report.

References

[1] Fan Chung. *Spectral Graph Theory.* American Mathematical Society, 1997.

[2] T-H. Hubert Chan, Anand Louis, Zhihao Gavin Tang, Chenzi Zhang. *Spectral Properties of Hypergraph Laplacian and Approximation Algorithms.* J. ACM, 65(3):15:1–15:48, 2018.

[3] T-H. Hubert Chan, Zhibin Liang. *Generalizing the Hypergraph Laplacian via a Diffusion Process with Mediators.* arXiv:1804.11128v1 [cs.DM] 30 Apr 2018.

[4] Yuichi Yoshida. *Cheeger Inequalities for Submodular Transformations.* arXiv:1708.08781v2 [cs.DS] 3 Apr 2018.

[5] Tasuku Soma, Yuichi Yoshida. *Spectral Sparsification of Hypergraphs.* arXiv:1807.04974v1 [cs.DS] 13 Jul 2018.

[6] C. Zhang, S. Hu, Z. G. Tang, and T-H. Hubert Chan. *Re-revisiting learning on hypergraphs: Confidence interval and subgradient method.* In ICML, volume 70 of Proceedings of Machine Learning Re- search, pages 4026–4034. PMLR, 2017.