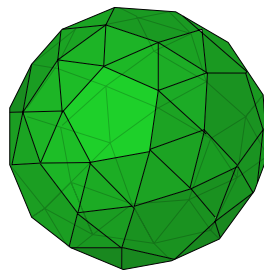


Projects in Mathematics and Applications

SPECTRAL CLUSTERING

Ngày 7 tháng 8 năm 2021

Vũ Phan Thăng Long ^{*} [†]Nguyễn Bá Khôi Nguyễn
Võ Minh Quân [‡] [§]Vòng Vĩnh Toàn



^{*}Đại học École Polytechnique

[†]Đại học Khoa học Tự nhiên TP.HCM

[‡]Trường THPT chuyên Hùng Vương

[§]Trường THPT chuyên Hùng Vương

Lời cảm ơn

Trước hết, chúng tôi xin gửi lời cảm ơn trân trọng đến với ban tổ chức trại hè Toán học và Ứng dụng PiMA, và các đơn vị tài trợ đã tạo điều kiện tốt nhất cho chúng tôi trong suốt quãng thời gian hai tuần tham gia trại. Mặc dù là năm đầu tiên tổ chức online do các nguyên nhân khách quan, nhưng các anh chị trong ban tổ chức đã cố gắng hết mình để trại hè có thể diễn ra thành công và suôn sẻ. Những kiến thức mà chúng tôi tích lũy được trong 2 tuần tại trại thực sự có ý nghĩa rất lớn đối với mỗi thành viên của nhóm, và là hành trang quý báu cho bản thân mỗi bạn trên con đường theo đuổi khoa học sau này.

Chúng tôi cũng gửi lòng biết ơn chân thành nhất đến các anh chị mentor đã nhiệt tình chỉ dạy và dẫn dắt chúng tôi trên những bước đầu trong công việc nghiên cứu. Chúng tôi cũng muốn gửi lời cảm ơn đến các giảng viên khách mời đã mang đến cho tất cả các bạn trại sinh những kinh nghiệm thiết thực về vấn đề nghiên cứu cũng như cơ hội việc làm sau này.

Cuối cùng xin cảm ơn tất cả các bạn trại sinh đã cùng nhau đồng hành với chúng mình trong kỳ trại vừa qua.

Do thời gian học tập cũng như nghiên cứu có hạn nên những thiếu sót là không thể tránh khỏi. Vì vậy nhóm chúng em rất mong nhận được sự đóng góp của các anh chị mentor cũng như các bạn để đề tài có thể được hoàn chỉnh hơn.

Tóm tắt nội dung

Bài báo cáo chủ yếu xoay quanh ý nghĩa toán học của thuật toán Unnormalized spectral clustering, các phiên bản cải tiến của nó và các vấn đề liên quan như Similarity matrix và Eigengap heuristic, đồng thời đưa ra một vài ví dụ cụ thể để minh họa cách vận hành các phương pháp này.

Mục lục

1	Đặt vấn đề	1
2	Cơ sở lý thuyết	1
3	Unnormalized spectral clustering	1
3.1	Xử lý dữ liệu đầu vào	1
3.2	Mô hình hóa bài toán	1
3.3	Thuật toán unnormalized spectral clustering	1
3.4	Eigengap heuristic	1
4	Áp dụng mô hình	1
5	Các phiên bản cải tiến	1
5.1	Symmetric Laplacian matrix	1
5.2	Random walk Laplacian matrix	1
6	Kết luận	1

1 Đặt vấn đề

2 Cơ sở lý thuyết

3 Unnormalized spectral clustering

3.1 Xử lý dữ liệu đầu vào

3.2 Mô hình hóa bài toán

3.3 Thuật toán unnormalized spectral clustering

3.4 Eigengap heuristic

4 Áp dụng mô hình

5 Các phiên bản cải tiến

5.1 Symmetric Laplacian matrix

5.2 Random walk Laplacian matrix

6 Kết luận

Tài liệu

- [1] David Benson-Putnins, Magaret Bonfardin, Meagan E. Magnoni, and Daniel Martin. Spectral clustering and visualization: A novel clustering of Fischer’s Iris dataset.
- [2] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2013.
- [3] Bojan Mohar. Some applications of laplace eigenvalues of graphs. pages 225–275, 1997.
- [4] Jianbo Shi and Jitendra Malik. Normalized cut and image segmentation. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 22(8), 2000.
- [5] Ulrike von Luxburg. A tutorial on spectral clustering. 2007.