CS 860 Spectral graph theory, Spring 2019, Waterloo.

Lecture 1: Introduction

We will start with the course information, and then an overview of the technical content of the course, and finally a review of some linear algebra for the course.

Course information

Course homepage: https://cs.uwaterloo.ca/~lapchi/cs860

In this course, we will learn how to use eigenvalues and eigenvectors to design algorithms and prove theorems, and more broadly to use linear algebraic and continuous techniques to study combinatorial problems.

This is a research oriented course. Our focus will be on doing active research in spectral graph theory.

Course requirement: 40%, class participation and homework
60%, course project.

You are expected to participate in class discussions, including asking and answering questions in lectures, discussing course material and open problems in piazza, and discussing homework solutions in piazza.

There will be two to three homeworks in total.

The main component is the course project, which includes two presentations and a written report.

In the project, the focus is on finding interesting problems, reading the relevant literature, and making some original progress to the problem.

The first presentation is the project proposal, explaining why you find the problem interesting and describing the related literature and your plan. This will be a week around midterm.

The second presentation and the written report will be about the progress that you have made. This will be at the end of the term.

Prerequisties: good background in linear algebra, probability, calculus, algorithms, discrete mathematics.

This course is mathematically oriented, with lots of proofs and calculations.

Make sure to allocate enough time if you decide to take this course as it may be quite demanding.

References : Notes will be provided. No textbooks.

Take a look at the course notes for CS 798 in Fall 2015 to get a good idea about the content.

Course overview

There are some recent breakthroughs in using ideas and techniques from spectral graph theory to design algorithms and prove new theorems for combinatorial problems.

In this course, we aim to study these new techniques and new connections, and try to make some progress in pushing them further.

Spectral graph theory is not a new topic. In the 80s, an important connection was made between the second largest eigenvalue of the adjacency matrix and the "expansion" of a graph.

This connection made by Cheeger's inequality is useful in different areas.

- Construction of expander graphs (sparse and well-connected), with various applications in TCS.
- Analysis of mixing time of random walks, useful in random sampling and approximate counting.
- Graph partitioning, the spectral partitioning method is a popular heuristic in practice.

Recently, there are a few generalizations of Cheeger's inequality using the k-th largest eigenvalue, e.g. to partition the graph into k parts using the first k eigenvectors.

We will study these in the beginning of the course.

Then, we study random walks on graphs, the connection between second eigenvalue and mixing time.

We will also see how to use random walks to design "local" algorithms for graph partitioning.

We will study "expander graphs" in details, the graphs on which random walks mix quickly.

Then, we will study a very recent active research area about "high-dimensional expanders".

This is used to analyze random walks on "simplicial complex" (a set system closed under taking subsets) and to answer a long standing open problem in random sampling called the matroid expansion conjecture (including sampling random spanning trees as special case).

This is an exciting new component in this edition of the course.

This is roughly the first (larger) half of the course.

After this, we will have the first project presentations.

Through the study of random walks, we will come across the idea of interpreting the graph as an electrical network, which is used to analyse hitting times of random walks.
This idea has found surprising applications recently. One direction is to use these concepts for graph sparsification, where the spectral perspective proved to be the right way to look at this.
Another direction is to use these concepts in combinatorial / convex optimization, which we won't have time to discuss.

Finally, we study an exciting new technique called the method of interlacing polynomials.
This is a new probabilistic method showing the existence of good combinatorial objects, but the proofs involve some mathematical concepts such as real-stable polynomials.
This method is used to make some breakthroughs in constructing expander graphs (Ramanujan graphs) and partitioning into expander graphs (the Kadison-Singer problem).
We will also discuss an amazing application of these ideas in designing approximation algorithms for the asymmetric traveling salesman problem.
The ideas developed in graph sparsification (e.g. barrier argument) is also a key component in this last part of the course.

There are some important topics that are not covered, such as eigenvalues of random matrices.
You can take a look at the project page for references on a list of interesting topics.

Eigenvalues and eigenvectors

Given an $n \times n$ matrix A, a nonzero vector v is an eigenvector of A if $Av = \lambda v$ for some scalar λ, which is called an eigenvalue associated with the eigenvector v.
The set of eigenvalues of A is given by the set of solutions to $\det(A - \lambda I) = 0$, the characteristic polynomial.
For an λ with $\det(A-\lambda I)=0$, any vector $v \neq 0$ in the kernel/nullspace of $A - \lambda I$ is an associated eigenvector.

Real symmetric matrices

Our starting point of spectral graph theory is the following spectral theorem for real symmetric matrices.

Theorem. Let A be an $n \times n$ real symmetric matrix. Then there is an orthonormal basis of $\mathbb{R}^n$

consisting of eigenvectors of A, and the corresponding eigenvalues are real numbers.

We will include a proof of this theorem in next section.

The above theorem applies to the adjacency matrices of undirected graphs, but not for directed graphs.

This is the main reason that spectral graph theory is much more developed for undirected graphs.

It has been an open direction to develop spectral graph theory for directed graphs, hypergraphs, etc.

Let $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ be the orthonormal basis of eigenvectors guaranteed by the above spectral theorem, with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Let V be the $n \times n$ matrix with the i-th column being v_n, i.e. $V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}$.

Let D be the $n \times n$ diagonal matrix with the (i,i)-th entry being λ_i, i.e. $D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$.

Then the conditions $Av_i = \lambda_i v_i \;\; \forall 1 \le i \le n$ can be compactly written as $AV = VD$.

Since the columns in V form an orthonormal basis, we have $V^T V = I$ and thus $V^{-1} = V^T$.

So, we can rewrite $AV = VD$ as $A = VDV^{-1} = VDV^T = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} - v_1 - \\ \vdots \\ - v_n - \end{bmatrix}$.

This representation is very convenient in computations.

Powers of matrices

To compute A^k, we observe that it is just $A^k = (VDV^T)^k = (VDV^T)(VDV^T)\cdots(VDV^T) = VD^kV^T$ as $V^TV = I$.

Since D is a diagonal matrix, D^k is easy to compute, i.e. $D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$, $D^k = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}$.

This is very useful, in analyzing random walks, as P^t is the transition matrix of the random walk after t steps where P is the transition matrix in one step.

We will use the eigenvalues of the transition matrix to bound the mixing time later.

Eigen-decomposition

Since $v_1, \ldots, v_n$ form an orthonormal basis, any $x \in \mathbb{R}^n$ can be written as $c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$.

By orthonormality, $\langle x, v_i \rangle = \langle c_1 v_1 + \cdots + c_n v_n, v_i \rangle = c_1 \langle v_1, v_i \rangle + \ldots + c_i \langle v_i, v_i \rangle + \ldots + c_n \langle v_n, v_i \rangle = c_i$.

Therefore, $x = \langle x, v_1 \rangle v_1 + \langle x, v_2 \rangle v_2 + \ldots + \langle x, v_n \rangle v_n$

$$= v_1 v_1^T x + v_2 v_2^T x + \ldots + v_n v_n^T x$$

$$= \left(v_1 v_1^T + v_2 v_2^T + \ldots + v_n v_n^T \right) x ..$$

This is true for all x, and hence $v_1v_1^T + v_2v_2^T + \dots + v_nv_n^T = I$.

Multiplying both sides by A, we get

$$Ax = A(v_1v_1^T + v_2v_2^T + \dots + v_nv_n^T)x$$
$$= (\lambda_1 v_1v_1^T + \lambda_2 v_2v_2^T + \dots + \lambda_n v_nv_n^T)x.$$

Thus, $A = \lambda_1 v_1v_1^T + \dots + \lambda_n v_nv_n^T$.

Finally, we claim that $A^{-1} = \frac{1}{\lambda_1}v_1v_1^T + \frac{1}{\lambda_2}v_2v_2^T + \dots + \frac{1}{\lambda_n}v_nv_n^T$ if $\lambda_i \neq 0$ for all i.

because $(\lambda_1 v_1v_1^T + \dots + \lambda_n v_nv_n^T)(\frac{1}{\lambda_1}v_1v_1^T + \dots + \frac{1}{\lambda_n}v_nv_n^T) = v_1v_1^T + \dots + v_nv_n^T = I$.

Later on, we will use this idea to define the "pseudo-inverse" of a matrix A, when A is not of full rank.

Positive semidefinite matrices

This is an important definition, an analog of a matrix being non-negative.

Fact Let A be a real symmetric matrix. The following statements are equivalent.

1. A is positive semidefinite, i.e. all eigenvalues of A are non-negative.
2. For any $x \in \mathbb{R}^n$, we have $x^TAx \geq 0$, i.e. all quadratic forms are non-negative.
3. $A = U^TU$ for some matrix $U \in \mathbb{R}^{n\times n}$.

proof Recall that a real symmetric matrix A can be written as VDV^T.

(1) $\Rightarrow$ (3) Since all eigenvalues are non-negative, we can write $A = (VD^{\frac{1}{2}})(D^{\frac{1}{2}}V^T)$ where $D^{\frac{1}{2}}$ is the $n\times n$ diagonal matrix with $D_{i,i} = \sqrt{\lambda_i}$.

Therefore, by letting $U = VD^{\frac{1}{2}}$, we see that A can be written as UU^T.

(3) $\Rightarrow$ (2) $x^TAx = x^TUU^Tx = \langle U^Tx, U^Tx \rangle = \|U^Tx\|_2^2 \geq 0$ for any $x \in \mathbb{R}^n$.

(2) $\Rightarrow$ (1) We prove the contrapositive, $\neg(1) \Rightarrow \neg(2)$.

If v is an eigenvector with negative eigenvalue, then $v^TAv = \lambda v^Tv = \lambda\|v\|_2^2 < 0$. □

We will use the notation $A \succcurlyeq 0$ for A is a positive semidefinite matrix.

This is the basic of "semidefinite programming", a powerful generalization of linear programming.

Unfortunately, we will not be able to see it in this course.

Trace = sum of eigenvalues

Fact Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A. Then, $\sum_{i=1}^n \lambda_i = \text{trace}(A)$ where trace of

A is defined as the sum of diagonal entries of A.

Proof Consider $\det(\lambda I - A)$.

Its roots are the eigenvalues of A, and so $\det(\lambda I - A) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$.

Note that the coefficients of $\lambda^{n-1} = -\sum_{i=1}^{n} \lambda_i$.

On the other hand, $\det(\lambda I - A) = \det \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{22} & \lambda - a_{22} & & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix}$.

By the (expansion) definition of the determinant, the coefficient of λ^{n-1} only appears in the term $(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$, which is $-\sum_{i=1}^{n} a_{ii}$.

Therefore, $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii} = \text{trace}(A)$. □

Exercise: Can you express $\det(A)$ using its eigenvalues?

Spectral theorem for real symmetric matrices (follow the presentation of [GR])

Theorem Let A be a real symmetric $n \times n$ matrix. Then there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A, and the corresponding eigenvalues are real numbers.

First we prove that there is a real eigenvalue.

Claim 1 Let A be a real symmetric matrix.

If u and v are eigenvectors with different eigenvalues, then u and v are orthogonal.

Proof Let $Au = \alpha u$ and $Av = \beta v$ where $\alpha \neq \beta$.

Then $v^T A u = v^T (Au) = v^T \alpha u = \alpha \langle v, u \rangle$.

On the other hand, $v^T A u = (v^T A) u = \beta v^T u = \beta \langle v, u \rangle$, because A is symmetric.

Now $\alpha \langle v, u \rangle = \beta \langle v, u \rangle$ and $\alpha \neq \beta$ implies that $\langle u, v \rangle = 0$, as required. □

Claim 2 The eigenvalues of a real symmetric matrix, if exist, are real numbers.

Proof Let $Au = \lambda u$. By taking the complex conjugate, we get $A\bar{u} = \bar{\lambda}\bar{u}$.

So $\bar{u}$ is also an eigenvector of A.

Since $u^T \bar{u} > 0$, by Claim 1, λ and $\bar{\lambda}$ cannot be of distinct values, and so λ is real. □

Since $\det(A - \lambda I) = 0$ always has a solution, Claim 2 implies that there is a real eigenvalue.

Now we would like to use induction to finish the proof.

To do this, we consider the vectors orthogonal to the existing eigenvectors, and show that there is an eigenvector in that subset of vectors.

We say a subspace U is <u>A-invariant</u> if $Au \in U$ for every $u \in U$, e.g. if U is the subspace generated by eigenvectors then U is A-invariant.

<u>Claim 3</u> Let A be a real symmetric matrix. If U is an A-invariant subspace, then $U^\perp$ is also A-invariant, where $U^\perp = \{ v \mid \langle v,u \rangle = 0 \ \forall u \in U \}$ is the set of vectors orthogonal to the vectors in U.

<u>Proof</u> Let $v \in U^\perp$ and $u \in U$. Then $v^T A u = v^T (Au) = v^T u' = 0$ where $u' \in U$.

It also means that $(v^T A) u = 0$ for all $v \in U^\perp$ and $u \in U$

So $v^T A$ is in $U^\perp$ for all $v \in U^\perp$.

Since A is symmetric, it implies that $Av \in U^\perp$ for all $v \in U^\perp$.

Hence $U^\perp$ is also A-invariant, as required. $\square$

<u>Claim 4</u> Let A be a real symmetric matrix. If U is a nonzero A-invariant subspace, then U contains a real eigenvector of A.

<u>Proof</u> Let R be a matrix whose columns form an orthonormal basis of U.

Then $AR = RB$ for some square matrix B, since each column of AR is a linear combination of the columns of R.

Then $R^T A R = B$, which implies that B is symmetric.

Therefore, there exists u such that $Bu = \lambda u$.

This implies that $ARu = RBu = \lambda Ru$.

Note that $Ru \neq 0$ for $u \neq 0$ because the columns of R are linearly independent.

Thus $Ru \in U$ is an eigenvector contained in U. $\square$

Now we can finish the proof by induction.

Let $\{u_1, \dots, u_m\}$ be an orthonormal set of eigenvectors of A, where $1 \le m < n$.

Let U be the subspace that they generate. Then U is A-invariant.

By Claim 3 $U^\perp$ is also A-invariant.

By claim 4, $U^{\perp}$ contains an eigenvector, and hence repeating this argument will finish the proof.

References [GR] Algebraic graph theory, by Godsil and Royle.