# CS 798: Algorithmic Spectral Graph Theory

Lap Chi Lau

Fall 2015

## A Brief Forward

Compiling Professor Lau's handwritten notes on CS798 in LaTeX.

- Patrick Wong

# Contents

# Lecture 1: Course Introduction

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

## 1.1 Course Information

Objective: learn how to use eigenvalues and eigenvectors to design algorithms and prove theorems. More generally, we will use linear algebraic and continuous techniques to solve combinatorial problems.

Course page: https://cs.uwaterloo.ca/~lapchi/cs798

Course Evaluation:
- Homework 50%, 2-3 assignments
- Project 50%

Project:
- To study one topic in-depth
- Some suggested topics in the project page (also served as further references for the course)
- encourage to choose a topic close to your research interest
- usually survey type, should be related to spectral grpah theory

Prerequisites: linear algebra, probability, algorithms, discrete mathematics (graphs)

All of them are covered in undergraduate computer science curriculum

Could get a good idea about the background of linear algebra in today's review

Should like math, as there will be lots of proofs and calculations

May be demanding, as definitions and techniques building up (e.g. missing two weeks will be tough)

Pace may be fast. Material may be too much

References: notes will be provided. No textbook.

You can see the lecture notes in my previous offering in 2012 to get a good idea, but there are new topics and some presentations of old topics will be changed.

You can also see the notes of similar courses in other universities as linked in the course page. I usually provide more details in the notes than in the lectures.

## 1.2 Course Overview

There are some recent breakthroughs in using ideas and techniques from spectral graph theory to design better algorithms for graph problems (both better in approximation and faster running time), and also prove new theorems. In this course, we aim to study these new techniques and see the new connections.

Spectral graph theory is not a new topic.

About thirty years ago, an important connection was made between the second eigenvalue and graph expansion. This connection made by Cheeger's inequality is useful in different areas:

- construction of expander graphs, which have various applications in theory and practice
- analysis of mixing time in random walks, with many applications in sampling and counting
- graph partitioning, and the spectral partitioning algorithm is a widely used heuristic in practice

We will study these in the beginning of the course.

Most of the early results in spectral graph theory are about the second eigenvalue.
In the last few years, researchers have found new connections between higher eigenvalues and graph properties, including the use of the k-th eigenvalue to find a small non-expanding set, to partition the graph into k non-expanding sets, and to do better analysis of the spectral partitioning algorithm and SDP-based approximation algorithms. We will discuss some of these, but probably not all details.

Then, we will study random walks on graphs, and see the classical connection between the second eigenvalue and the mixing time.
After that, we see some recent results in using random walks to find a small non-expanding set. This problem is closely related to the unique games conjecture, which is about the limits of polynomial time approximation algorithms. Combining ideas from random walks and higher eigenvalues gives the best attempt to go beyond the unique games barrier.
This is about half the course. We will also talk about constructions of expander graphs and small-set expanders if time permits.

Through the study of random walks, we will come across the idea of interpreting the graph as an electrical network, which is used to analyze hitting times of random walks. This has found surprising applications recently.
One line of research is to use ideas from convex optimization to compute these electrical flows quickly, and then somehow these could be used to compute electrical flows faster! The other direction is to use these concepts for graph sparsification, where the spectral perspective proved to be the right way to look at this combinatorial problem.

Finally, we study an exciting new technique called the method of interlacing polynomials. This is a new probablistic method showing the existence of good combinatorial objects, but the proofs involve some mathematical concepts like 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 the last part of the course.

This overview is still very sketchy. In class, we will elaborate with more precise definition and more details. This is an unusual exception that the lecture has more details than the notes.

Some important topics that will not be covered include SDP-based approximation algorithms, eigenvalues of random graphs, and applications of the spectral methods in machine learning.

The topics are theoretically-oriented. I will try to focus on the underlying techniques that are relevant in broader contexts.

If you are interested in learning more about these topics before the lecture (e.g. previewing the method of interlacing polynomials), the course project page provides further references.

## 1.3 Background on Linear Algebra

### 1.3.1 Linear transformation, linear independence, nullspace, rank

Let $A \in R^{mxn}$ be a $n$ $x$ $n$ matrix, and $x \in R^n$ be a column vector (we use $x$ to denote a column vector and $x^T$ to denote a row vector).
One can view $A$ as a matrix or linear transformation from n-dimensional vectors to n-dimensional vectors.
A set of vectors $v_1, v_2, ..., v_k$ are linearly independent if $c_1v_1 + c_2v_2 + ... + c_kv_k = 0$ implies $c_1 = c_2 = ... = c_k = 0$; otherwise they are linearly dependent.
The nullspace of $A$ (or the kernel of $A$) is defined as $nullspace(A) = \{x \in R^n \mid Ax = 0\}$ and its dimension is defined as $null(A)$.
The range of $A$ (or the image of $A$) is defined as $range(A) = \{Ax \mid x \in R^n\}$ and its dimension is defined as the rank of A, denoted as $rank(A)$.
It is known that $rank(A) + null(A) =$ if $A$ is a $n$ $x$ $n$ matrix.
It follows from the definition that $rank(A)$ is the number linearly independent columns of $A$.

### 1.3.2 determinant

The determinant of $A$, denoted by $det(A)$, is defined recursively as

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot det(\tilde{A}_{ij})$$

where $a_{ij}$ is the (i, j)-th entry of $A$ and $\tilde{A}$ is the matrix obtained from $A$ by deleting the $i$-th row and $j$-th column.
We can unfold the recursion and get an equivalent (expansion) definition:

$$det(A) = \sum_{\sigma \in S^n} sgn(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

where $\sigma : [n] \to [n]$ is a permutation (or a bijection) of the indices, and $sgn(\sigma) = +1$ if $\sigma$ is even and $sgn(\sigma) = -1$ if $\sigma$ is odd. More precisely, $sgn(\sigma) = (-1)^{inv(\sigma)}$ where $inv(\sigma) = |\{(i, j) \mid i < j \text{ and } \sigma(i) > \sigma(j)\}|$ is the number of inversions of $\sigma$

A basic property of the determinant is the following:

$$det\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u+kv \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = det\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} + k \cdot det\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix}$$

which follows from the expansion definition of the determinant.
From this it follows that

$$det\begin{pmatrix} \vdots \\ ca_i \\ \vdots \end{pmatrix} = c{\cdot}det\begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix},\ det\begin{pmatrix} v \\ v \\ \vdots \end{pmatrix} = 0,\ det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = -det\begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix},\ det\begin{pmatrix} \vdots \\ a_i + ca_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}$$

Therefore, we can compute the determinant by reducing $A$ into an upper triangular matrix by elementary row operations, and prove that $det(A) \neq 0$ if and only if $rank(A) = n$ (because $rank(A) = n$ if and only if it can be reduced to an upper triangular matrix where all the diagonal entries are nonzero).

One can also prove that $det(AB) = det(A)det(B)$, as this is true for elementary matrices and one can write $A$ and $B$ as products of elementary matrices if $rank(A) = rank(B) = n$

### 1.3.3 Eigenvalues and Eigenvectors

A nonzero vector $x$ is an eigenvector of $A$ if there exists $\lambda$ such that $Ax = \lambda x$, and the scalar $\lambda$ is called an eigenvalue of $A$.

Note that $Ax = \lambda x$ iff $(A - \lambda I)x = 0$ where $I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ is the identity matrix.

There is a nonzero vector x satisfying $(A - \lambda I)x = 0$

$$\begin{aligned} &\iff nullspace(A - \lambda I) \neq \{0\} \\ &\iff rank(A - \lambda I) < n \\ &\iff det(A - \lambda I) = 0, i.e.\ det\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21}-\lambda & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}-\lambda & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix} \end{aligned}$$

By the (expansion) definition of the determinant, $det(A - \lambda I)$ is a polynomial of $\lambda$ of degree $n$.
This is called the characteristic polynomial of $A$.
In the last offering of this course, the characteristic polynomials did not play a major role, but they will be very important in the last part of this offering as they will be the key object in the interlacing method. Also, determinants will play an important role in the last part.

One can compute the characteristic polynomial by Gaussian elimination or interpolation.
Any root of this polynomial is an eigenvalue, and any vector in $nullspace(A - \lambda I)$ is an eigenvector.
Geometrically, an eigenvector is a direction that is "fixed" (but can be scaled) by the linear transformation.
So not all matrices have an eigenvector with a real eigenvalue, e.g. the rotation matrix has no direction fixed, and thus no eigenvector with a real eigenvalue.
But we will see that all real symmetric matrices have real eigenvalues, a corollary of the spectral theorem that we will prove.

### 1.3.4 Orthogonality

To state the spectral theorem, we need one more concept.

### 1.3.5 Spectral theorem for real symmetric matrices

## 1.4 Applications of the spectral theorem

### 1.4.1 Power of matrices

# Lecture 2: Graph Spectrum

forward

## 2.1 Adjacency Matrix

this