| | |
|---|---|
| **Spectral Graph Theory** | Lecture 4 |

# Bounding Eigenvalues

*Daniel A. Spielman* September 10, 2012

## 4.1 About these notes

These notes are not necessarily an accurate representation of what happened in class. The notes written before class say what I think I should say. The notes written after class way what I wish I said.

All statements in these notes that can be made mathematically rigorous should be taken with a grain of salt and a shot of Tequila.

## 4.2 Overview

It is unusual when one can actually explicitly determine the eigenvalues of a graph. Usually one is only able to prove loose bounds on some eigenvalues. In this lecture, I will introduce two important techniques for proving such bounds. The first is the Courant-Fischer Theorem, which provides a more powerful characterization of eigenvalues as solutions to optimization problems than the one we derived before. This theorem is useful for doing things like proving upper bounds on the largest eigenvalue of a matrix.

The other technique we will use is one which I call "Graphic Inequalities". It allows one to compare one graph with another, and prove things like lower bounds on the largest eigenvalue of a matrix.

## 4.3 The Courant-Fischer Theorem

I gave a hint of the Courant-Fischer Theorem earlier in the lecture. I'll do the rest of it now.

**Theorem 4.3.1** (Courant-Fischer Theorem)**.** *Let $A$ be a symmetric matrix with eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$. Then,*

$$\mu_k = \max_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S)=k}} \min_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \min_{\substack{T \subseteq \mathbb{R}^n \\ \dim(T)=n-k+1}} \max_{\boldsymbol{x} \in T} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

For example, consier the case $k = 1$. In this case, $S$ is just the span of $\boldsymbol{v}_1$ and $T$ is all of $\mathbb{R}^n$. For general $k$, the optima will be achieved when $S$ is the span of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ and when $T$ is the span of $\boldsymbol{v}_k, \ldots, \boldsymbol{v}_n$.

*Proof.* We will just verify the first characterization of $\mu_k$. The other is similar.

First, let's verify that $\mu_k$ is achievable. Let $S_k$ be the span of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$. For every $\boldsymbol{x} \in S_k$, we can write

$$\boldsymbol{x} = \sum_{i=1}^{k} c_i \boldsymbol{v}_i,$$

so,

$$\frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{\sum_{i=1}^{k} \mu_i c_i^2}{\sum_{i=1}^{k} c_i^2} \geq \frac{\sum_{i=1}^{k} \mu_k c_i^2}{\sum_{i=1}^{k} c_i^2} = \mu_k.$$

To verify that this is in fact the maximum, let $T_k$ be the span of $\boldsymbol{v}_k, \ldots, \boldsymbol{v}_n$. As $T_k$ has dimension $n-k+1$, for any $S$ of dimension $k$ the intersection of $S$ with $T_k$ has dimension at least 1. So,

$$\min_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \leq \min_{\boldsymbol{x} \in S \cap T_k} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

Any such $\boldsymbol{x}$ may be expressed as

$$\boldsymbol{x} = \sum_{i=k}^{n} c_i \boldsymbol{v}_i,$$

and so

$$\frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{\sum_{i=k}^{n} \mu_i c_i^2}{c_i^2} \leq \frac{\sum_{i=k}^{n} \mu_k c_i^2}{c_i^2} = \mu_k.$$

We conclude that for all subspaces $S$ of dimension $k$,

$$\min_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \leq \mu_k.$$

□

## 4.4 Bounds on $\lambda_2$

The Courant-Fischer Theorem provides a simple way of proving upper bounds on $\lambda_2$–the second-smallest eigenvalue of the Laplacian. Recall

$$\lambda_2 = \min_{\boldsymbol{v}: \boldsymbol{v}^T \mathbf{1} = 0} \frac{\boldsymbol{v}^T L \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}}.$$

So, every vector $\boldsymbol{v}$ orthogonal to $\mathbf{1}$ provides an upper bound on $\lambda_2$:

$$\lambda_2 \leq \frac{\boldsymbol{v}^T L \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}}.$$

When we use a vector $\boldsymbol{v}$ in this way, we call it a *test vector*.

The Courant-Fischer theorem is not as helpful when we want to prove lower bounds on $\lambda_2$. To prove lower bounds, we need the form with a maximum on the outside, which gives

$$\lambda_2 \geq \max_{S:\dim(S)=n-1} \min_{\boldsymbol{v}\in S} \frac{\boldsymbol{v}^T L \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}}.$$

This is not too helpful, as it is difficult to prove lower bounds on

$$\min_{\boldsymbol{v}\in S} \frac{\boldsymbol{v}^T L \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}}$$

over a space $S$ of large dimension. So, we need a new technique.

## 4.5 Graphic Inequalities

I begin by recalling an extremely useful piece of notation that is used in the Optimization community. For a symmetric matrix $A$, we write

$$A \succcurlyeq 0$$

if $A$ is positive semidefinite. That is, if

$$\boldsymbol{v}^T A \boldsymbol{v} \geq 0,$$

for all $\boldsymbol{v}$. We similarly write

$$A \succcurlyeq B$$

if

$$\boldsymbol{v}^T A \boldsymbol{v} \geq \boldsymbol{v}^T B \boldsymbol{v}$$

for all $\boldsymbol{v}$. This is the same as

$$A - B \succcurlyeq 0.$$

The relation $\preccurlyeq$ is an example of a *partial order*. It applies to some pairs of symmetric matrices, while others are incomparable. But, for all pairs to which it does apply, it acts like an order. For example, we have

$$A \succcurlyeq B \text{ and } B \succcurlyeq C \text{ implies } A \succcurlyeq C,$$

and

$$A \succcurlyeq B \text{ implies } A + C \succcurlyeq B + C,$$

for symmetric matrices $A$, $B$ and $C$.

I find it convenient to overload this notation by defining it for graphs as well. Thus, I'll write

$$G \succcurlyeq H$$

if $L_G \succcurlyeq L_H$. For example, if $G = (V, E)$ is a graph and $H = (V, F)$ is a subgraph of $G$, then

$$L_G \succcurlyeq L_H.$$

To see this, recall that the Laplacian of a graph can be expressed as the sum of the Laplacians of its edges. As $F \subseteq E$, we get

$$L_G = \sum_{e \in E} L_e = \sum_{e \in F} L_e + \sum_{e \in E - F} L_e \succcurlyeq \sum_{e \in F} L_e = L_H,$$

as

$$\sum_{e \in E - F} L_e \succcurlyeq 0.$$

In this proof, I have used the notation $L_e$ to indicate the Laplacian consisting of the graph containing just the edge $e$.

This notation is most powerful when we consider some multiple of a graph. Thus, I could write

$$G \succcurlyeq c \cdot H,$$

for some $c > 0$. What is $c \cdot H$? It is the same graph as $H$, but the weight of every edge is multiplied by $c$.

Using the Courant-Fischer Theorem, we can prove

**Lemma 4.5.1.** *If $G$ and $H$ are graphs such that*

$$G \succcurlyeq c \cdot H,$$

*then*

$$\lambda_k(G) \geq c\lambda_k(H),$$

*for all $k$.*

*Proof.* The Courant-Fischer Theorem tells us that

$$\lambda_k(G) = \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = k}} \max_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T L_G \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \geq c \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = k}} \max_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T L_H \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = c\lambda_k(H).$$

□

**Corollary 4.5.2.** *Let $G$ be a graph and let $H$ be obtained by either adding an edge to $G$ or increasing the weight of an edge in $G$. Then, for all $i$*

$$\lambda_i(G) \leq \lambda_i(H).$$

This lemma provides an easy way of bounding how much the eigenvalues of a graph can change if we change the weights on some of its edges.

**Lemma 4.5.3.** *Let $G = (V, E, w)$ and $H = (V, E, z)$ be two weighted graphs that differ only in their edge weights. Then*

$$G \succcurlyeq \min_{e \in E} \frac{w(e)}{z(e)} H.$$

*Proof.* Recall that the Laplacian of a graph may be expressed as the sum of the Laplacians of its edges. So,

$$L_G = \sum_{e \in E} w(e) L_e = \sum_{e \in E} \frac{w(e)}{z(e)} z(e) L_e \geq \left( \min_{e \in E} \frac{w(e)}{z(e)} \right) \sum_{e \in E} z(e) L_e = \left( \min_{e \in E} \frac{w(e)}{z(e)} \right) L_H.$$

□

## 4.6 Approximations of Graphs

An idea that we will use in later lectures is that one graph approximates another if their Laplacian quadratic forms are similar. For example, we will say that $H$ is a $c$-approximation of $G$ if

$$cH \succcurlyeq G \succcurlyeq H/c.$$

Surprising approximations exist. For example, expander graphs are very sparse approximations of the complete graph. For example, the following is known.

**Theorem 4.6.1.** *For every $\epsilon > 0$, there exists a $d > 0$ such that for all sufficiently large $n$ there is a $d$-regular graph $G_n$ that is a $(1+\epsilon)$-approximation of $K_n$.*

These graphs have many fewer edges than the complete graphs!

In a later lecture we will also prove that every graph can be well-approximated by a sparse graph.

## 4.7 The Path Inequality

By now you should be wondering, "how do we prove that $G \succcurlyeq c \cdot H$ for some graph $G$ and $H$?" Not too many ways are known. We'll do it by proving some inequalities of this form for some of the simplest graphs, and then extending them to more general graphs. For example, we will prove

$$(n-1) \cdot P_n \succcurlyeq G_{1,n}. \tag{4.1}$$

That is, $n-1$ times the path of length $n-1$ from vertex 1 to $n$ is greater than the edge from 1 to $n$.

The following very simple proof of this inequality was discovered by Sam Daitch.

**Lemma 4.7.1.**

$$(n-1) \cdot P_n \succcurlyeq G_{1,n}.$$

*Proof.* We need to show that for every $\boldsymbol{x} \in \mathbb{R}^n$,

$$(n-1) \sum_{i=1}^{n-1} (\boldsymbol{x}(i+1) - \boldsymbol{x}(i))^2 \geq (\boldsymbol{x}(n) - \boldsymbol{x}(1))^2.$$

For $1 \leq i \leq n-1$, set

$$\boldsymbol{\delta}(i) = \boldsymbol{x}(i+1) - \boldsymbol{x}(i).$$

The inequality we need to prove then becomes

$$(n-1)\sum_{i=1}^{n-1} \boldsymbol{\delta}(i)^2 \geq \left(\sum_{i=1}^{n-1} \boldsymbol{\delta}(i)\right)^2.$$

But, this is just the Cauchy-Schwartz inequality. I'll remind you that Cauchy-Schwartz just follows from the fact that the inner product of two vectors is at most the product of their norms:

$$(n-1)\sum_{i=1}^{n-1} \boldsymbol{\delta}(i)^2 = \|\mathbf{1}_{n-1}\|^2 \|\boldsymbol{\delta}\|^2 = (\|\mathbf{1}_{n-1}\| \|\boldsymbol{\delta}\|)^2 \geq \left(\mathbf{1}_{n-1}^T \boldsymbol{\delta}\right)^2 = \left(\sum_{i=1}^{n-1} \boldsymbol{\delta}(i)\right)^2.$$

□

While I won't cover it in lecture, I will also state the version of this inequality for weighted paths.

**Lemma 4.7.2.** *Let $w_1, \ldots, w_{n-1}$ be positive. Then*

$$G_{1,n} \preccurlyeq \left(\sum_{i=1}^{n-1} \frac{1}{w_i}\right) \sum_{i=1}^{n-1} w_i G_{i,i+1}.$$

*Proof.* Let $\boldsymbol{x} \in \mathbb{R}^n$ and set $\boldsymbol{\delta}(i)$ as in the proof of the previous lemma. Now, set

$$\boldsymbol{\gamma}(i) = \boldsymbol{\delta}(i)\sqrt{w_i}.$$

Let $\boldsymbol{w}^{-1/2}$ denote the vector for which

$$\boldsymbol{w}^{-1/2}(i) = \frac{1}{\sqrt{w_i}}.$$

Then,

$$\sum_i \boldsymbol{\delta}(i) = \boldsymbol{\gamma}^T \boldsymbol{w}^{-1/2},$$

$$\left\|\boldsymbol{w}^{-1/2}\right\|^2 = \sum_i \frac{1}{w_i},$$

and

$$\|\boldsymbol{\gamma}\|^2 = \sum_i \boldsymbol{\delta}(i)^2 w_i.$$

So,

$$\begin{aligned}\boldsymbol{x}^T L_{G_{1,n}} \boldsymbol{x} &= \left(\sum_i \boldsymbol{\delta}(i)\right)^2 = \left(\boldsymbol{\gamma}^T \boldsymbol{w}^{-1/2}\right)^2 \\ &\leq \left(\|\boldsymbol{\gamma}\| \left\|\boldsymbol{w}^{-1/2}\right\|\right)^2 = \left(\sum_i \frac{1}{w_i}\right) \sum_i \boldsymbol{\delta}(i)^2 w_i = \left(\sum_i \frac{1}{w_i}\right) \boldsymbol{x}^T \left(\sum_{i=1}^{n-1} w_i L_{G_{i,i+1}}\right) \boldsymbol{x}.\end{aligned}$$

□

### 4.7.1 Bounding $\lambda_2$ of a Path Graph

I'll now demonstrate how to use these techniques to bound the second-smallest eigenvalue of a path graph. We will see that it is approximately $c/n^2$, for some constant $c$. In the next lecture, I will show you exactly what the eigenvalues of the path graph are.

First, let's use a test vector to get an upper bound. Consider the vector $\boldsymbol{x}$ such that $\boldsymbol{x}(i) = (n+1) - 2i$, for $1 \leq i \leq n$. This vector satisfies $\boldsymbol{x} \perp \mathbf{1}$, so

$$\begin{aligned}
\lambda_2(P_n) &\leq \frac{\sum_{1 \leq i < n} (x(i) - x(i+1))^2}{\sum_i x(i)^2} \\
&= \frac{\sum_{1 \leq i < n} 2^2}{\sum_i (n+1-2i)^2} \\
&= \frac{4(n-1)}{(n+1)n(n-1)/3} && \text{(clearly, the denominator is } n^3/c \text{ for some } c\text{)} \\
&= \frac{12}{n(n+1)}.
\end{aligned}$$

So, we can easily obtain an upper bound on $\lambda(P_n)$ that is of the right order of magnitude.

To prove a lower bound on $\lambda_2(P_n)$, we will prove that some multiple of the path is at least the complete graph. To this end, recall that

$$L_{K_n} = \sum_{i<j} L_{G_{i,j}},$$

and that

$$\lambda_2(K_n) = n.$$

For every edge $(i,j)$ in the complete graph, we apply the only inequality available in the path:

$$G_{i,j} \preccurlyeq (j-i) \sum_{k=i}^{j-1} G_{k,k+1} \preccurlyeq (j-i) P_n.$$

Summing this inequality over all edges $(i,j) \in K_n$ gives

$$K_n = \sum_{i<j} G_{i,j} \preccurlyeq \sum_{i<j} (j-i) P_n.$$

To finish the proof, we compute

$$\sum_{1 \leq i < j \leq n} (j-i) = \sum_{k=1}^{n-1} k(n-k) = n(n+1)(n-1)/6.$$

So,

$$\frac{n(n+1)(n-1)}{6} \cdot L_{P_n} \succcurlyeq L_{K_n}.$$

Applying Lemma 4.5.1, we obtain

$$\lambda_2(P_n) \geq \frac{6}{(n+1)(n-1)}.$$

This only differs from our lower bound by a factor of 2.

### 4.7.2 The Complete Binary Tree

Let's do the same analysis with the complete binary tree.

One way of understanding the complete binary tree of depth $d+1$ is to identify the vertices of the tree with strings over $\{0,1\}$ of length at most $d$. The root of the tree is the empty string. Every other node has one ancestor, which is obtained by removing the last character of its string, and two children, which are obtained by appending one character to its label.

Alternatively, you can describe it as the graph on $n = 2^{d+1} - 1$ nodes with edges of the form $(i, 2i)$ and $(i, 2i+1)$ for $i < n$. We will name this graph $T^d$. Pictures of this graph appear below.

Pictorially, these graphs look like this:

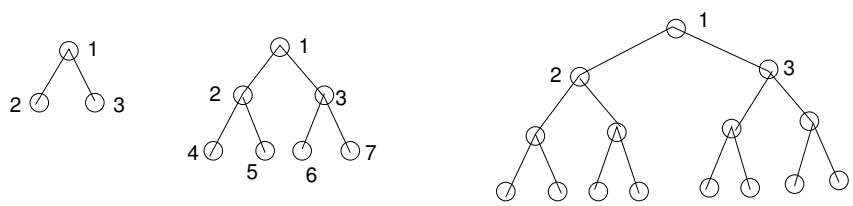


Figure 4.1: $T_1$, $T_2$ and $T_3$. Node 1 is at the top, 2 and 3 are its children. Some other nodes have been labeled as well.

Let's first upper bound $\lambda_2(T_d)$ by constructing a test vector $x$. Set $x(1) = 0$, $x(2) = 1$, and $x(3) = -1$. Then, for every vertex $u$ that we can reach from node 2 without going through node 1, we set $x(u) = 1$. For all the other nodes, we set $x(u) = -1$.

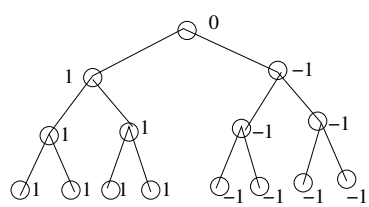


Figure 4.2: The test vector we use to upper bound $\lambda_2(T_3)$.

We then have

$$\lambda_2 \leq \frac{\sum_{(i,j)\in T_n}(x_i - x_j)^2}{\sum_i x_i^2} = \frac{(x_1 - x_2)^2 + (x_1 - x_3)^2}{n-1} = 2/(n-1).$$

We will again prove a lower bound by comparing $T_n$ to the complete graph. For each edge $(i,j) \in K_n$, let $T_n(i,j)$ denote the unique path in $T$ from $i$ to $j$. This path will have length at most $2d$. So, we have

$$K_n = \sum_{i<j} G_{i,j} \preccurlyeq \sum_{i<j} (2d) T_n(i,j) \preccurlyeq \sum_{i<j} (2\log_2 n) T_n = \binom{n}{2} (2\log_2 n) T_n.$$

So, we obtain the bound

$$\binom{n}{2}(2\log_2 n)\lambda_2(T_n) \geq n,$$

which implies

$$\lambda_2(T_d) \geq \frac{1}{(n-1)\log_2 n}.$$

In the problem set, I will ask you to improve this lower bound to $1/cn$ for some constant $c$.