Spectral Graph Theory	Instructor: Padraic Bartlett

Lecture 2: Spectral Theory and Decomposition Problems, part 2/2

Week 4 — *Mathcamp 2011*

In our last lecture, we introduced the **Lagrangian**, a function $\mathbb{R}^n \to \mathbb{R}$ defined as follows:

$$f_G(\mathbf{v}) = \langle A_G\mathbf{v}, \mathbf{v}\rangle = (A_G\mathbf{v})^T \cdot \mathbf{v} = \sum_{\{i,j\}\in E(G)} 2v_iv_j.$$

We introduced this function because we wanted something that would allow us to distinguish between graphs that have "many" positive eigenvalues and graphs that have "many" negative eigenvalues. This function did this, in the following sense:

Proposition 1 *The Lagrangian function f_G is (positive-semidefinite/positive-definite/negative-semidefinite/negative-definite) on the space generated by all of the eigenvectors corresponding to (nonnegative/positive/nonpositive/negative, respectively) eigenvalues.*

To illustrate the usefulness of this tool, here's an application:

Theorem 2 *The complete graph K_n cannot be decomposed into $\leq n-2$ complete bipartite graphs.*

Proof. First, notice that (given the intuition we've developed earlier) we would expect this to be a problem: on one hand, we have $\text{Spec}(K_n) = \{(n-1)^1, (-1)^{n-1}\}$, which has a lot of negative eigenvalues, while $\text{Spec}(K_{m,n}) = \{\pm\sqrt{mn}, 0^{n-2}\}$ has a lot of nonnegative eigenvalues.

But how does our tool tell us this? In other words, suppose that we had some such decomposition of K_n into $n-2$ complete bipartite graphs $G_1, \ldots G_{n-2}$; to make our lives easier, add vertices to each of these G_i's so that they're all on n vertices. Because adding an unconnected vertex to a graph changes the spectrum by adding a 0-eigenvalue, these graphs all have spectrum $\{\pm ab, 0^{n-2}\}$.

So: how can we use the Lagrangian? Well, one nice observation we can make about the Lagrangian is that it distributes across graph decompositions: i.e. because $A_{K_n} = A_{G_1} + \ldots A_{G_{n-2}}$, we have

$$f_{K_n} = \sum_{i=1}^{n-2} f_{G_i}.$$

How can we use this? Well, notice that each A_{G_i} has a $n-1$-dimensional space corresponding to its nonnegative eigenvectors – call it U_i, say. Then, because we have $n-2$ of these spaces, the intersection of all of these spaces is at least dimension 2. But this means that there is a 2-dimensional space on which the function $\sum_{i=1}^{n-2} f_{G_i}$ is positive semidefinite.

However: this is actually f_{K_n} in disguise! And f_{K_n} is negative-definite on a space of dimension $n-1$; i.e. there's only one dimension of space in which f_{K_n} doesn't take on negative values! So this is clearly a contradiction.

Excellent! The Lagrangian works! Win.

...what else can we do with it?

1 Finding Structure in Graphs

In the above argument, we used the Lagrangian to find problems with given graph decompositions. Can we reverse this kind of argument – i.e. can we use the Lagrangian to show that graphs with certain properties **must** contain certain kinds of subgraphs?

It turns out that the answer is yes! Specifically, we can use the Lagrangian to pick out the **clique number** of a graph, in the following way: let

$$f(G) = \max_{\mathbf{s} \in S} f_G(\mathbf{s}),$$

where $S = \{\mathbf{s} : \sum_{i=1}^{n} s_i = 1, s_i \geq 0\}$. In a certain sense, then, $f(G)$ is picking out the pockets of "density" in our graph G: i.e. if you think of $\mathbf{s}$ as a weighting of the vertices on our graph, the maximal value of $f_G(\mathbf{s})$ is attained where $\mathbf{s}$ is concentrated on the vertices with "lots" of edges.

We make this rigorous with the following two observations:

Proposition 3 *$f(K_n) = (n-1)/n$.*

Proof. Simply note that

$$\begin{aligned}
f(G) &= \max_{\mathbf{s} \in S} f_G(\mathbf{s}) \\
&= \max_{\mathbf{s} \in S} \sum_{\{i,j\} \in E} 2a_{ij} s_i s_j \\
&= \max_{\mathbf{s} \in S} \sum_{i=1}^{n} \sum_{j \neq i} s_i s_j \\
&= \max_{\mathbf{s} \in S} \sum_{i=1}^{n} s_i \sum_{j \neq i} s_j \\
&= \max_{\mathbf{s} \in S} \sum_{i=1}^{n} s_i (1 - s_i) \\
&= \max_{\mathbf{s} \in S} \left(\sum_{i=1}^{n} s_i \right) - \left(\sum_{i=1}^{n} s_i^2 \right) \\
&= \max_{\mathbf{s} \in S} 1 - \left(\sum_{i=1}^{n} s_i^2 \right),
\end{aligned}$$

which is maximized by the point in S closest to the origin: i.e. $(1/n, 1/n, \ldots 1/n)$. At this point, we have $f_G(1/n, \ldots, 1/n) = (n-1)/n$, as claimed.

Proposition 4 *Let G be a graph with clique number k. Then $f(G) = (k-1)/k$.*

Proof. Let S' be the collection of points $\mathbf{s} \in S$ where f_G attains its maximum, and amongst these points let $\mathbf{y}$ be a point in this collection with the **smallest support**: i.e. one with the most coördinates equal to 0 in our collection.

As it turns out, if you do this and look at the coördinates of $\mathbf{y}$ that aren't zero, you get a complete graph! To see why, suppose not: i.e. that you have two coördinates y_1, y_2 with

- $y_1, y_2 > 0$, and
- $\{y_1, y_2\} \notin E(G)$.

Let

$$y_1 \sum_{j=1}^{n} 2a_{1j}y_j$$

be the portion of $f_G(\mathbf{y})$ that depends on y_1, and

$$y_2 \sum_{j=1}^{n} 2a_{2j}y_j$$

be the portion of $f_G(\mathbf{y})$ that depends on y_2. One of the two sums $\sum_{j=1}^{n} 2a_{1j}y_j, \sum_{j=1}^{n} 2a_{2j}y_j$ has to be larger (or equal); assume it's the first, without any loss of generality. Then, because there isn't an edge between y_1 and y_2, we have that the two inequalities

$$\begin{aligned} y_2 \sum_{j=1}^{n} 2a_{2j}y_j &\leq y_2 \sum_{j=1}^{n} 2a_{1j}y_j, \\ y_1 \sum_{j=1}^{n} 2a_{1j}y_j &= y_1 \sum_{j \neq 2} 2a_{1j}y_j \end{aligned}$$

hold, which forces

$$\Rightarrow \qquad y_1 \sum_{j=1}^{n} 2a_{1j}y_j + y_2 \sum_{j=1}^{n} 2a_{2j}y_j \leq (y_1 + y_2) \sum_{j=1}^{n} 2a_{1j}y_j$$

and thus that $f_G(\mathbf{y}) \leq f_G(y_1 + y_2, 0, y_3, \ldots y_n)$. This element is also in S and has one more zero-coördinate than $\mathbf{y}$ did: a contradiction to $\mathbf{y}$'s minimality! So we've proven our claim.

So: we've used the Lagrangian to not just show that certain decompositions are impossible, but also that certain graphs must have various structures – i.e. $f(G)$ is completely determined by the clique number! This allows us to prove the following result with almost no effort at all:

Theorem 5 *Suppose that r is a constant and G is a graph with n vertices and m edges, with $m > \frac{r-1}{2r}n^2$. Then G contains K_r as a subgraph.*

Proof. If you plug in $(1/n, \ldots 1/n)$ into f_G, you get

$$\frac{2m}{n^2} \geq \frac{r-1}{r}.$$

But, on the other hand, you know that

$$(\omega(G) - 1)/(\omega(G)) = f(G) \geq f_G(1/n, \ldots 1/n),$$

and thus that $\omega(G) \geq r$.

Proving this theorem without linear algebraic machinery is a lot of extra work; here, we get it for essentially free!