Notes for Lecture 9

Last time we have defined the *edge expansion* of a graph G by

$$\begin{aligned} h \; &:= \; \min_{\substack{S \subseteq V \\ |S| \le |V|/2}} \frac{\text{edges}(S, V-S)}{|S|}, \quad \text{where} \\ \text{edges}(A,B) \; &:= \; \sum_{u \in A} \sum_{v \in B} (\#\text{edges between } u \text{ and } v). \end{aligned}$$

We also defined the *adjacency matrix* M of G by

$$M(u,v) := \#\text{edges between } u \text{ and } v.$$

Note that G may be a multigraph, i.e., the number of edges between two vertices may exceed 1. Being a symmetric matrix, M has a real spectrum which can be ordered as $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$, with the corresponding orthonormal system of eigenvectors $x_1, x_2, \ldots, x_n$. We also proved that, for a d-regular graph, $\lambda_1 = d$, $x_1 = \mathbf{1}/\sqrt{n}$, where $\mathbf{1} := (1,1,\ldots,1)$ and $n := |V|$, and that $h = 0$ if and only if $\lambda_1 = \lambda_2$.

The goal of this lecture is to prove the following theorem.

Theorem 1 *For a d-regular graph,*

1. $h \ge (d - \lambda_2)/2$,
2. $h \le \sqrt{2d(d-\lambda_2)}$.

PROOF: In the previous lecture, we have proved that

$$\lambda_2 = \max_{x \perp \mathbf{1}} \frac{xMx^T}{xx^T}.$$

Consider the quadratic form $\sum_{u,v} M(u,v)(x(u)-x(v))^2$. We can rewrite it as

$$\sum_{u,v} M(u,v)(x(u)-x(v))^2 = 2d \sum_v x^2(v) - 2 \sum_{u,v} M(u,v) x(u) x(v) = 2dxx^T - 2xMx^T. \tag{1}$$

Therefore

$$\lambda_2 = \max_{x \perp \mathbf{1}} \frac{2dxx^T - \sum_{u,v} M(u,v)(x(u)-x(v))^2}{2xx^T} = d - \min_{x \perp \mathbf{1}} \frac{\sum_{u,v} M(u,v)(x(u)-x(v))^2}{2xx^T}.$$

To prove the first statement of the theorem, take a set S such that $h = \text{edges}(S, V-S)/|S|$ and $|S| \le |V|/2$. Let $p := |S|/n$, $q := 1 - p = |V-S|/n$, and define

$$x(v) := \begin{cases} q & v \in S, \\ -p & v \notin S \end{cases}$$

For this vector x, we get $xx^T = npq^2+nqp^2 = npq(p+q) = npq$, whereas $x \perp \mathbf{1}$, and $\sum_{u,v} M(u,v)(x(u)-x(v))^2 = 2hnp(p+q)^2 = 2hnp$. Hence

$$d - \lambda_2 \leq \frac{2hnp}{2npq} = \frac{h}{q} \leq 2h.$$

This proves the first statement.

To prove the second statement, we will establish three auxiliary facts.

Claim 1. For all $y \in \mathbb{R}^n$,

$$\sum_{u,v} M(u,v)|y^2(u) - y^2(v)| \leq \sqrt{2dyy^T - 2yMy^T}\sqrt{4dyy^T}.$$

A proof of Claim 1 follows from the Cauchy-Schwarz inequality and from formula (1):

$$\begin{aligned}\sum_{u,v} M(u,v)|y^2(u) - y^2(v)| &= \sum_{u,v} M(u,v)|y(u) - y(v)| \cdot |y(u) + y(v)| \\ &\leq \sqrt{\sum_{u,v} M(u,v)(y(u) - y(v))^2}\sqrt{\sum_{u,v} M(u,v)(y(u) + y(v))^2} \\ &\leq \sqrt{\sum_{u,v} M(u,v)(y(u) - y(v))^2}\sqrt{\sum_{u,v} 2M(u,v)(y^2(u) + y^2(v))} \\ &= \sqrt{2dyy^T - 2yMy^T}\sqrt{4dyy^T}.\end{aligned}$$

Claim 2. Suppose x is an eigenvector corresponding to eigenvalue λ_2: $xM = \lambda_2 x$ and such that $\#\{v : x(v) > 0\} \leq n/2$ (the latter can always be achieved by replacing x by $-x$ if necessary). Define a vector y by $y(v) := \max\{x(v), 0\}$. Then $yM \geq \lambda_2 y$ componentwise.

Indeed, if $x(v) \geq 0$, then $(yM)(v) \geq (xM)(v) = \lambda_2 x(v) = \lambda_2 y(v)$. If $x(v) < 0$, then $y(v) = 0$, but $(yM)(v) \geq 0$.

Claim 3. For the vector y defined in Claim 2,

$$\sum_{u,v} M(u,v)|y^2(u) - y^2(v)| \geq 2hyy^T.$$

Here is a proof of Claim 3: Let us arrange the components of y in nonincreasing order

$$y(v_1) \geq y(v_2) \geq \cdots \geq y(v_n).$$

Suppose t of these components are strictly positive: $y(v_t) > y(v_{t+1}) = \cdots = y(v_n) = 0$. Denote by K the set where the jumps occur: $K := \{k : y(v_k) > y(v_{k+1})\}$. With this notation, rewrite the sum $\sum_{u,v} M(u,v)|y^2(u) - y^2(v)|$ as

$$2\sum_{i=1}^{t}\sum_{j=i+1}^{n} M(v_i,v_j)(y^2(v_i) - y^2(v_j)) = 2\sum_{k\in K}\sum_{i\leq k}\sum_{j>k} M(v_i,v_j)(y^2(v_k) - y^2(v_{k+1})).$$

The last equality comes from the fact that in case there are several elements of the set K between two indices i and j, the sum $\sum_{k\in K, i\le k<j}(y^2(v_k) - y^2(v_{k+1}))$ telescopes to $y^2(v_i) - y^2(v_j)$.

Now, for each $k = 1, \ldots, n$, denote by L_k the set $\{v_i : i \le k\}$; for $k = 0$, set $L_0 := \emptyset$. With this notation, we have $\sum_{i\le k}\sum_{j>k} M(v_i, v_j) \ge h|L_k|$, so

$$2\sum_{k\in K}\sum_{i\le k}\sum_{j>k} M(v_i, v_j)(y^2(v_k) - y^2(v_{k+1})) \ge 2\sum_{k\in K} h|L_k|(y^2(v_k) - y^2(v_{k+1})).$$

Using summation by parts and denoting by k' the element of K preceding k, the last expression can be written as

$$2h\sum_{k\in K}(|L_k| - |L_{k'}|)y^2(v_k) = 2h\sum_{k\in K}\#\{v : y(v) = y(v_k)\}y^2(v_k) = 2h\sum_{v} y^2(v) = 2hyy^T.$$

This proves Claim 3.

Finally, we can combine the three results we just proved. First not that the vector y from Claim 2 is not a zero vector, since x is a nonzero vector orthogonal to the vector $\mathbf{1}$, hence x has nonzero entries of both signs. Therefore,

$$\begin{aligned} h &\le \frac{\sum_{u,v} M(u,v)|y^2(u) - y^2(v)|}{2yy^T} \le \frac{\sqrt{2dyy^T - 2yMy^T}\sqrt{4dyy^T}}{2yy^T} \\ &\le \frac{\sqrt{2dyy^T - 2\lambda_2 yy^T}\sqrt{4dyy^T}}{2yy^T} = \sqrt{2d(d-\lambda_2)}. \end{aligned}$$

This finishes the proof. $\square$