Chapter 7
Second neighbor representation and the Regularity Lemma

Nachdiplomvorlesung by László Lovász
ETH Zürich, Spring 2014

Perhaps the easiest way of assigning a vector to each node is to use the corresponding column of the adjacency matrix. Even this easy construction has some applications, but (surprisingly) a more interesting geometric representation can be obtained by using the columns of the *square* of the adjacency matrix. This construction is related to the Regularity Lemma of Szemerédi [4], one of the most important tools of graph theory.

1 Regularity partitions

1.1 Matrix norms

For a matrix $M \in \mathbb{R}^{U\times V}$ and for $S \subseteq U,\ T \subseteq V$, let $A(S,T) = \sum i \in S, j \in TA(S,T)$. The *cut-norm* of a matrix $M \in \mathbb{R}^{m\times n}$:

$$\|M\|_\square = \max_{S\subseteq[m],T\subseteq[n]} |M(S,T)|.$$

We also need the (entrywise) 1-norm and max-norm of a matrix:

$$\|M\|_1 = \sum_{i=1}^m \sum_{j=1}^n |M_{i,j}|, \qquad \|M\|_\infty = \max_{i,j} |M_{i,j}|.$$

Clearly $\|M\|_\square \le \|M\|_1 \le mn\|M\|_\infty$. The cut-norm satisfies

$$|\mathbf{x}^\mathsf{T} M\mathbf{y}| \le 4\|M\|_\square \|\mathbf{x}\|_\infty \|\mathbf{y}\|_\infty \tag{1}$$

for any to vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$. If one of the vectors $\mathbf{x}$ or $\mathbf{y}$ is nonnegative, then we can replace the coefficient 4 by 2; if both are nonnegative, we can replace it by 1.

If $n = 1$, then the cut norm is within a factor of 2 of the 1-norm, but in general these norms may be very far apart. But we do have the following reverse inequality.

Lemma 1.1 *For any two matrices $M \in \mathbb{R}^{n\times m}$ and $A \in \{0,1\}^{m\times k}$, we have*

$$\|MA\|_1 \le 4m\|M\|_\square \|A\|_\infty.$$

If $A \ge 0$, then the coefficient 4 can be replaced by 2.

Proof. Fix any $j \in V$, and let s_i be the sign of $(MA)_{ij}$, $\mathbf{s} = (s_1, \dots s_m)^\mathsf{T}$. Then by (1),

$$\sum_i |(MA)_{ij}| = \sum_i s_i (MA)_{ij} = \mathbf{s}^\mathsf{T} M(A\mathbf{e}_j) \le 4\|M\|_\Box \|A\|_\infty,$$

since $\|\mathbf{s}\|_\infty \le 1$ and $\|A\mathbf{e}_j\|_\infty \le \|A\|_\infty$. Summing over j, the first assertion of the lemma follows. The second assertion follows similarly. $\Box$

For several other useful properties of the cut norm, we refer to [2], Section 8.1.

1.2 Regularity Lemmas

Very briefly we collect some facts about regularity lemmas.

Let $G = (V, E)$ be a graph on n nodes. Let $e_G(V_i, V_j)$ denote the number of edges in G connecting V_i and V_j. If G is edge-weighted, then the weights of edges should be added up. If $i = j$, then we define $e_G(V_i, V_i) = 2|E(G[V_i])|$.

Let $\mathcal{P} = \{V_1, \dots, V_k\}$, be a partition of V. We define the weighted graph $G_\mathcal{P}$ on V by taking the complete graph and weighting its edge uv by $e_G(V_i, V_j)/(|V_i|\,|V_j|)$ if $u \in V_i$ and $v \in V_j$. We allow the case $u = v$ (and then, of course, $i = j$), so G has loops. Let A be the adjacency matrix of G, then we denote by $A_\mathcal{P}$ the weighted adjacency matrix of $G_\mathcal{P}$. The matrix $A_\mathcal{P}$ is obtained by averaging the entries of A over every block $V_i \times V_j$:

$$(A_\mathcal{P})_{u,v} = \frac{A(V_i, V_j)}{|V_i|\,|V_j|} \qquad \text{for } u \in V_i, v \in V_j.$$

It will be convenient to define, for two graphs G and H on the same set of n nodes,

$$d_\Box(G, H) = \frac{1}{n^2}\|A_G - A_H\|_\Box.$$

The Regularity Lemma says, roughly speaking, that the node set of every graph has an equitable partition $\mathcal{P}$ into a “small” number of classes such that $G_\mathcal{P}$ is “close” to G. Various (non-equivalent) forms of this lemma can be proved, depending on what we mean by “close”. The version we need was proved by Frieze and Kannan [1].

Lemma 1.2 (Weak Regularity Lemma) *For every $k \ge 1$ and every graph $G = (V, E)$, V has a partition $\mathcal{P}$ into k classes such that*

$$d_\Box(G, G_\mathcal{P}) \le \frac{2}{\sqrt{\log k}}.$$

We do not require here that $\mathcal{P}$ be an equitable partition; it is not hard to see that this version implies that there is also an equitable partition with similar property, just we have to double the error bound.

There is another version of the weak regularity lemma, that is sometimes even more useful.

Lemma 1.3 *For every matrix $A \in \mathbb{R}^{n\times n}$ and every $k \ge 1$ there are k 0-1 matrices $Q_1, \dots, Q_k$ of rank 1 and k real numbers a_i such that $\sum_i a_i^2 \le 4$ and*

$$\Big\| A - \sum_{i=1}^{k} a_i Q_i \Big\|_\square \le \frac{4n^2}{\sqrt{k}}.$$

Note that a 0-1 matrix Q of rank 1 is determined by two sets $S, T \subseteq [n]$ by the formula $Q_{u,v} = \mathbb{1}(u \in S)\mathbb{1}(v \in T)$.

2 Second neighbors

We define the *2-neighborhood representation* of a graph $G = (V, E)$ as the map $i \mapsto \mathbf{u}_i$, where $\mathbf{u}_i$ is the column of A_G^2 corresponding to node i. Squaring the matrix seems unnatural, but it is crucial. We define a distance between the nodes, called the 2-*neighborhood distance* (sometimes the *similarity distance*), by

$$d(s,t) = \frac{1}{n^2}\|\mathbf{u}_s - \mathbf{u}_t\|_1.$$

A set of nodes $S \subseteq V$ is an *ε-cover*, if for every point $v \in V$ $d(v, S) \le \varepsilon$.(Here, as usual, $d(v, S) = \min_{s\in S} d(s, v)$.) We say that S is an *ε-packing*, if $d(u, v) \ge \varepsilon$ for every $u, v \in S$. An *ε-net* is a set that is both an ε-packing and an ε-cover. It is clear that a maximal ε-packing must be an ε-covering (and so, and ε-net). If S is an ε-cover, then no (2ε)-packing can have more than $|S|$ elements. On the other hand, every ε-cover has a subset that is both an ε-packing and a (2ε)-cover.

We also need the notion of an *average ε-cover*, which is a set $S \subseteq V$ such that $\sum_{v\in V} d(v, S) \le \varepsilon n$. An *average ε-net* is an average (2ε)-cover that is also an ε-packing. (It is useful to allow this relaxation by a factor of 2 here.) For every average ε-cover, a maximal subset that is an ε-packing is an average ε-net.

How to construct ε-nets? Assuming that we can compute $d(u, v)$ efficiently, we can go through the nodes in any order, and build up the ε-net S, by adding a new node v to S if $d(v, S) \le \varepsilon$. In other words, S is a greedily selected maximal ε-packing, and therefore an ε-net.

Average ε-nets also come in through a simple algorithm. Suppose that the underlying set $V(G)$ is too large to go through each node, but we have a mechanism to select a uniformly distributed random node (this is the setup in the theory of graph property testing). Again we build up S randomly, in each step generating a new random node and adding it to S if its distance from S is at least ε. We stop if for $1/\varepsilon^2$ consecutive steps no new has been added. The set S formed this way is an ε-packing, but not necessarily maximal. However, it is likely to be an average (2ε)-cover. Indeed, let t be the number of points x with $d(x, S) > \varepsilon$. These points are at distance at most 1 from S, and hence

$$\sum_{x\in V} d(x,S) \le (n-t)\varepsilon + t < n\varepsilon + t.$$

So if S is not an average (2ε)-cover, then $t > \varepsilon n$. The probability that we have not hit this set for $1/\varepsilon^2$ steps is less than $e^{-1/\varepsilon}$.

The following was proved (in a somewhat different form) in [3].

Theorem 2.1 *Let $G = (V,E)$ be a simple graph, and let d be its 2-neighborhood distance.*

(a) *If $S \subseteq V$ is an average ε-cover, then the Voronoi cells of S (with respect to the metric d) define a partition $\mathcal{P}$ of V such that $d_\square(G, G_\mathcal{P}) \le 8\sqrt{\varepsilon}$.*

(b) *If $\mathcal{P} = \{V_1,\dots,V_k\}$ is a partition of V such that $d_\square(G,G_\mathcal{P}) \le \varepsilon$, then we can select elements $s_i \in V_i$ so that $\{s_1,\dots,s_k\}$ is an average (4ε)-cover.*

Combining with the Weak Regularity Lemma, it follows that every graph has an average ε-cover, with respect to the 2-neighborhood metric, satisfying

$$|S| \le \exp\Big(O\Big(\frac{1}{\varepsilon^2}\Big)\Big).$$

It also follows that it has a similarly bounded average ε-net.

Proof. (a) Let $S = \{s_1,\dots,s_k\}$, and let $\mathcal{P} = \{V_1,\dots,V_k\}$ be the partition of V defined by the Voronoi cells of S (where $s_i \in V_i$). For $v \in S_t$, let $\phi(v) = s_t$. Define two matrices $P, Q \in \mathbb{R}^{V\times V}$ by

$$P_{ij} = \begin{cases} \dfrac{1}{|S_t|}, & \text{if } i,j \in S_t,\\ 0, & \text{otherwise.}\end{cases} \qquad Q_{ij} = \begin{cases} 1, & \text{if } j = \phi(i),\\ 0, & \text{otherwise.}\end{cases}$$

Clearly $P^2 = P$. If $\mathbf{x} \in \mathbb{R}^V$, then $P\mathbf{x}$ is the vector whose i-th coordinate is the average of $\mathbf{x}_j$ over the class of $\mathcal{P}$ containing i. Similarly, $PAP = A_\mathcal{P}$. The vector $Q^\mathsf{T}\mathbf{x}$ is obtained from $\mathbf{x}$ by replacing the entry in position $i \in S_t$ by the entry in position s_t.

Write $R = A - A_\mathcal{P} = A - PAP$, then $PRP = 0$. We want to show that $\|R\|_\square \le 8\sqrt{\varepsilon}$. It suffices to show that for every vector $\mathbf{x} \in \{0,1\}^V$,

$$\langle \mathbf{x}, R\mathbf{x}\rangle \le 2\sqrt{\varepsilon}\, n^2. \tag{2}$$

Let us write $\mathbf{y} = \mathbf{x} - P\mathbf{x}$, then $P\mathbf{y} = Q\mathbf{y} = 0$, $RP\mathbf{x} = 0$, and $A\mathbf{y} = R\mathbf{y}$. Hence

$$\begin{aligned}\langle \mathbf{x}, R\mathbf{x}\rangle &= \langle \mathbf{y}, R\mathbf{x}\rangle + \langle P\mathbf{x}, R\mathbf{x}\rangle = \langle \mathbf{x}, R\mathbf{y}\rangle + \langle P\mathbf{x}, R\mathbf{y}\rangle + \langle P\mathbf{x}, RP\mathbf{x}\rangle = \langle \mathbf{x} + P\mathbf{x}, A\mathbf{y}\rangle\\ &\le 2\|A\mathbf{y}\|_1 \le 2\sqrt{n}\|A\mathbf{y}\|_2. \end{aligned} \tag{3}$$

Here

$$\begin{aligned}\|A\mathbf{y}\|_2^2 = \langle \mathbf{y}, A^2\mathbf{y}\rangle = \langle \mathbf{y}, (A^2 - A^2Q)\mathbf{y}\rangle \le \|A^2 - A^2Q)\|_1 = \sum_k \|\mathbf{u}_k - \mathbf{u}_{\phi(k)}\|_1 \\ = \sum_k d(k, \phi(k))n^2 = \sum_k d(k, S)n^2 \le \varepsilon n^3.\end{aligned}$$

Combining with (3), this proves (2).

(b) Suppose that $\mathcal{P}$ is a weak regularity partition with error ε. Let $R = A - PAP$, then we know that $\|R\|_\square \le \varepsilon$. Let i, j be two nodes in the same partition class of $\mathcal{P}$. Then $Pe_i = Pe_j$, and hence $A(e_i - e_j) = R(e_i - e_j)$. Thus

$$\|\mathbf{u}_i - \mathbf{u}_j\|_1 = \|A^2(e_i - e_j)\|_1 = \|AR(e_i - e_j)\|_1 \le \|ARe_i\|_1 + \|ARe_j\|_1. \tag{4}$$

For every set $T \in \mathcal{P}$, choose a point $v_T \in T$ for which $\|ARe_i\|_1$ is minimal, and let $S = \{v_T :\ T \in \mathcal{P}\}$. Then using (4) and Lemma (1.1),

$$\begin{aligned}\sum_i d(i, S) \le \sum_{T\in\mathcal{P}} \sum_{i\in T} \|\mathbf{u}_i - \mathbf{u}_{v_T}\|_1 \le \sum_{T\in\mathcal{P}} \sum_{i\in T} \big(\|ARe_i\|_1 + \|ARe_j\|_1\big) \\ = \sum_i \|ARe_i\|_1 + \sum_{T\in\mathcal{P}} |T|\|ARe_{v_T}\|_1 \le 2\sum_i \|ARe_i\|_1 = 2n\|AR\|_1 \le 4\varepsilon n.\end{aligned}$$

This completes the proof. □

What can we say about ε-nets, not in the average but in the exact sense? Not much in a general metric space, since exceptional points, far from from everything else, must be included, and the number of these could be negligible relative to the whole size of the space, but not relative to the size of the average ε-net. It is somewhat surprising that in the case of graphs, even exact ε-nets can be bounded by a function of ε (although the bound is somewhat worse than for average ε-nets). The following result was proved by Alon (unpublished).

Theorem 2.2 *In any graph G, any set of k nodes contains two nodes s and t with*

$$d(s, t) \le 20\sqrt{\frac{\log\log k}{\log k}}.$$

In other words, every ε-net (or ε-packing) S in a graph, with respect to the 2-neighborhood metric, satisfies

$$|S| \le \exp\Big(O\Big(\frac{1}{\varepsilon^2}\log\frac{1}{\varepsilon}\Big)\Big).$$

Proof. Fix $s, t \in V$, then we have

$$d(s, t) = \frac{1}{n^2}\|A^2(\mathbf{e}_s - \mathbf{e}_t)\|_1.$$

Let us substitute for the second factor of A its decomposition in Lemma 1.3:

$$A = \sum_{i=1}^{q} a_i \mathbb{1}_{S_i} \mathbb{1}_{T_i}^{\mathsf{T}} + R, \tag{5}$$

where $q = \lfloor (\log k)/(\log\log k) \rfloor$, $\sum_i a_i^2 \le 4$, and $\|R\|_\square \le 4/\sqrt{q}$. We get

$$d(s,t) \le \frac{1}{n^2} \sum_{i=1}^{q} |a_i| \|\mathbb{1}_{S_i} \mathbb{1}_{T_i}^{\mathsf{T}} A(\mathbf{e}_s - \mathbf{e}_t)\|_1 + \|RA(\mathbf{e}_s - \mathbf{e}_t)\|_1.$$

The remainder term is small by the bound on R and by Lemma 1.1:

$$\|RA(\mathbf{e}_s - \mathbf{e}_t)\|_1 \le 4\|R\|_\square \le \frac{16}{\sqrt{q}}.$$

Considering a particular term of the first sum,

$$\|\mathbb{1}_{S_i}(\mathbb{1}_{T_i}^{\mathsf{T}} A\mathbf{e}_s - \mathbb{1}_{T_i}^{\mathsf{T}} A\mathbf{e}_t)\|_1 = \|\mathbb{1}_{S_i}\|_1 |\mathbb{1}_{T_i}^{\mathsf{T}} A\mathbf{e}_s - \mathbb{1}_{T_i}^{\mathsf{T}} A\mathbf{e}_t)| \le n|A(s,T_i) - A(t,T_i)|.$$

Hence

$$d(s,t) \le \frac{1}{n} \sum_{i=1}^{q} |a_i|\, |A(s,S_i) - A(t,S_i)| + \frac{16}{\sqrt{q}}.$$

Since $k > q^q$, by the pigeon-hole principle, any given set of k nodes contains two nodes s and t such that $|A(s,S_i) - A(t,S_i)| \le n/q$ for all i, and for this choice of s and t we have

$$d(s,t) \le \sum_{i=1}^{q} |a_i| \frac{1}{q} + \frac{16}{\sqrt{q}} \le \Big(\frac{1}{q}\sum_{i=1}^{q} |a_i|^2\Big)^{1/2} + \frac{16}{\sqrt{q}} \le \frac{18}{\sqrt{q}} < 20\sqrt{\frac{\log\log k}{\log k}}.$$

This proves the theorem. □

Example 2.3 Let $G = G(n,W)$ be obtained by selecting n random points on the d-dimensional unit sphere, and connecting two of these points x and y if $\measuredangle(x,y) \le 90°$. Then with high probability $d(x,y) \sim \sqrt{d}\measuredangle(x,y)$, and from this is follows that every ε-net in this graph has cardinality $\Omega(\sqrt{d}\varepsilon)^{-d}$. The best choice for d is $d \sim 1/\varepsilon^2$, which gives that every ε-net has at least $2^{\Omega(1/\varepsilon^2)}$ elements.

References

[1] A. Frieze and R. Kannan: Quick approximation to matrices and applications, *Combinatorica* **19** (1999), 175–220. .

[2] L. Lovász: *Large graphs, graph homomorphisms and graph limits*, Amer. Math. Soc., Providence, RI (2012).

[3] L. Lovász and B. Szegedy: Szemerédi's Lemma for the analyst, *Geom. Func. Anal.* **17** (2007), 252–270.

[4] E. Szemerédi: Regular partitions of graphs, *Colloque Inter. CNRS* (J.-C. Bermond, J.-C. Fournier, M. Las Vergnas and D. Sotteau, eds.) (1978) 399–401.