Le pseudo-aléa: objets et génération. Exercises

David Xiao <dxiao@liafa.jussieu.fr>
CNRS, LIAFA, Université Paris Didérot - Paris 7

March 7, 2012

1 Graphs and their spectra

Let $G = (V, E)$ be an undirected D-regular graph of size $n = |V|$ and let its normalized adjacency matrix be M, defined as $M_{i,j} = e(i, j)/D$ where $e(i, j)$ is the number of edges in G between vertices i and j (allowing for multiple edges). Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of M and let us suppose they are ordered so that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. Let $v_1, \ldots, v_n$ be the corresponding orthonormal eigenvectors.

1. Show that the eigenvalues of M lie in the interval $[-1, 1]$. Show that the uniform vector $u = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$ is an eigenvector of M with eigenvalue 1.
2. Show that if G has at least k connected components, then G has eigenvalue 1 with multiplicity at least k. (Stronger statement: In fact, the converse holds as well, and therefore the number of connected components equals the multiplicity of 1, but we will not prove this now.)
3. Let G^k denote the graph on the same vertex set V as G and where for all $i, j \in V$ the number of edges between i and j in G^k is the number of paths of length k between i, j in the original graph G (allowing for multiple edges between the same pair of points). Show that λ is an eigenvalue of G iff λ^k is an eigenvalue of G^k.
4. Show that if G is connected and bipartite then it has an eigenvalue of -1. (You may use the stronger statement of Item 2.)

2 Expander walk sampling and randomness-efficient error reduction

1. Fix G a (n, D, λ) expander. Fix any set $B \subseteq [n]$. Let $W = (W_0, \ldots, W_k)$ denote the steps of a random walk in G defined by picking $W_0 \leftarrow_{\mathrm{R}} [n]$ and letting W_i be a random neighbor of W_{i-1} for all $i \geq 1$. Let $\beta = |B|/n$ be the density of B in $[n]$. Prove the following:
   (a) Define the diagonal matrix P where the i'th diagonal is 1 if $i \in B$ and 0 otherwise. Prove that $\|PM\| \leq (\sqrt{\beta} + \lambda)$ (where $\|\cdot\|$ is the operator norm, *i.e.* $\|A\| = \max_{x \in \mathbb{R}^n} \|Ax\|_2/\|x\|_2$).
   (b) Let u denote the vector of the uniform distribution, $u = (1/n, \ldots, 1/n)^T$. Show that:
   $$\Pr[W_1, \ldots, W_k \in B] = |(PM)^k u|_1 \tag{2.1}$$
   (Notice we start from W_1, not W_0. This is a technicality that will simplify calculations later.)
   (c) Conclude that
   $$\Pr[W_1, \ldots, W_k \in B] \leq (\sqrt{\beta} + \lambda)^k \tag{2.2}$$
2. Fix a language L and an efficient algorithm A, such that for all $x \in \{0, 1\}^n$, A uses $m = \mathrm{poly}(n)$ random bits and satisfies:

$$\forall x \in L, \ \Pr[A(x; U_m) = 1] \geq 8/9$$
$$\forall x \notin L, \ \Pr[A(x; U_m) = 1] = 0$$

Namely, A is an efficient algorithm deciding L with one-sided error (only on positive instances). Suppose there exists a $(2^m, D, \lambda)$ expander with $D = O(1)$ and $\lambda < 1/6$.
For any k, construct an efficient algorithm A' that uses $m' = m + O(k)$ random bits such that $\forall x \in L,\ \Pr[A(x; U_{m'}) = 1] \geq 1 - 2^{-k}$ and $\forall x \notin L,\ \Pr[A'(x; U_{m'}) = 1] = 0$

3. Fix a language L and an efficient algorithm A, such that for all $x \in \{0,1\}^n$, A uses $m = \mathrm{poly}(n)$ random bits and satisfies:

$$\forall x \in \{0,1\}^n,\ \Pr[A(x; U_m) = L(x)] \geq 1 - 2^{-10}$$

Namely, A is an efficient algorithm deciding L with *two-sided* error. Suppose there exists a $(2^m, D, \lambda)$ expander with $D = O(1)$ and $\lambda < 2^{-5}$.
For any k, construct an efficient algorithm A' that uses $m' = m + O(k)$ random bits such that

$$\forall x \in \{0,1\}^n,\ \Pr[A(x; U_{m'}) = L(x)] \geq 1 - 2^{-k}$$

Hint: define A' using the majority of k samples taken by an expander walk, and to analyze the probability that A' errs, take a union bound over all possible subsets of steps of the walk $S \subseteq [k]$ with size $|S| \geq k/2$. Then, using a generalization of Equation 2.1, bound the probability that the steps of the walk in S are bad.

3 Binary error-correcting codes and ε-biased generators

Recall that we can naturally identify $\{0,1\}^n$ with the vector space $GF(2)^n$. Recall the following definitions:

Definition 3.1. $\mathcal{C} \subseteq \{0,1\}^n$ is a $[n,k,d]$ linear code if $\mathcal{C}$ is a linear subspace of $\{0,1\}^n$ with dimension k, and if for all distinct $x, y \in \mathcal{C}$ it holds that $|x-y|_H \geq d$ where $|\cdot|_H$ denotes the Hamming weight (number of non-zero entries) of a vector.

Definition 3.2. $G : \{0,1\}^s \to \{0,1\}^k$ is an ε-biased generator if for all linear functions $f : \{0,1\}^k \to \{0,1\}$, it holds that

$$|\Pr[f(G(U_s)) = 1] - \tfrac{1}{2}| \leq \varepsilon$$

Prove the following:

1. Given an ε-biased generator $G : \{0,1\}^s \to \{0,1\}^k$, one can construct a $[2^s, k, 2^s(\frac{1}{2} - \varepsilon)]$ linear code.
2. Is it possible to do the reverse, *i.e.* given $\mathcal{C}$ a $[n, k, n(\frac{1}{2} - \varepsilon)]$ linear code to construct an ε-biased generator? If so, give a construction. If not, explain why not.

4 Efficient constructions of combinatorial designs

Show that for any constant $K > 0$, one can find in time $\mathrm{poly}(m)$ a family of sets $S_1, \ldots, S_m \subseteq [9K \log m]$ with the following properties:

1. For all $i \in [m]$, $|S_i| = \sqrt{K} \log m$.
2. For all $i \neq j \in [m]$, $|S_i \cap S_j| \leq \log m$.

Hint: greedily build the family $S_1, \ldots, S_m$ one-by-one, and at each time $i < m$ prove that there exists a suitable S_{i+1} by using a probabilistic argument and the following version of the Hoeffding bound.

Lemma 4.1. *Fix any $T \subseteq [n]$. Suppose S is drawn as a random subset of size s out of $[n] = \{1, \ldots, n\}$. Then for all $\delta > 0$ the following holds:*

$$\Pr\left[|S \cap T| > (1+\delta)\tfrac{|T|}{n}\right] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{s|T|/n}$$