1 Shearer’s Lemma

Today we shall learn about Shearer’s Lemma, which is a generalization of the subadditivity of entropy.

**Lemma 1 (Shearer’s Lemma).** Let $X = X_1, \ldots, X_n$ be any random variables. If $S$ is any distribution on subsets of $\{1 \ldots n\}$, such for every $i$, $\Pr[i \in S] \geq \mu$, then $\mathbb{E}[H(X_S)] \geq \mu \cdot H(X)$.

(As an aside, we give a simple proof due to Jaikumar Radhakrishnan.)

**Proof** For $T = \{i_1, \ldots, i_k\}$ with $i_1 < i_2 < \cdots < i_k$, observe that

$$H(X_T) = H(X_{i_1}) + H(X_{i_2} | X_{i_1}) + \cdots + H(X_{i_k} | X_{i_{k-1}}, \ldots, X_{i_1})$$

$$\geq H(X_{i_1} | X_{<i_1}) + H(X_{i_2} | X_{<i_2}) + \cdots + H(X_{i_k} | X_{<i_k}),$$

where we used chain rule in the equality, and used the fact that entropy is only smaller if we condition on more variables, for the inequality.

Thus, we get that

$$\mathbb{E}_S[H(X_S)] \geq \mathbb{E}_S \left[ \sum_{i \in S} H(X_i | X_{<i}) \right]$$

$$= \sum_{i \in [n]} \Pr[i \in S] \cdot H(X_i | X_{<i}) \quad \text{whenever } i \text{ is not in } S, \text{ this term contributes 0}$$

$$\geq \mu \sum_{i \in [n]} H(X_i | X_{<i})$$

$$= \mu \cdot H(X)$$

2 Applications

Now, let’s start counting the number of cliques within a graph. We start with a simple example. Suppose $G = (V,E)$ is an undirected graph, $t$ is the number of triangles and $\ell$ is the number of edges.

**Proposition 2.** $t \leq (2\ell)^{3/2}/6$

**Proof** The proof is very similar to that of the triangles and vee problem we have seen. Let $X_1, X_2, X_3$ be uniformly random vertices forming a triangle. Then $H(X_1, X_2, X_3) = \log(6t)$, since each triangle can be written in 6 ways.

Let $S$ be a uniformly random subset of coordinates $\{1, 2, 3\}$ of size 2. Then for all $i$, $\Pr[i \in S] = 2/3$. By Shearer’s Lemma,

$$\mathbb{E}_S[H(X_S)] \geq \frac{2}{3} \log(6t),$$

*Based in part on lecture notes by Anup Rao and Jijiang Yan.*
so there exists $T \subset [1,2,3], |T| = 2$, for which $H(X_T) \geq \frac{2}{3} \log(6t)$. On the other hand $X_T$ is supported on edges of the graph, so $\log(2\ell) \geq H(X_T)$. This gives $2\ell \geq (6t)^{2/3}$, proving the bound. 

It is easy to see using a similar proof that if $a < b$ and $n_a$ is the number of cliques of size $a$ and $n_b$ is the number of cliques of size $b$, then $(b! \cdot n_b)^a \leq (a! \cdot n_a)^b$. Can we say something about arbitrary subgraphs (besides cliques)? It turns out that we can completely characterize the relationship between the number of subgraphs to the number of edges!

### 2.1 Counting Embeddings of Graphs

$N(T, \ell)$ is the maximum number of homomorphisms a graph $T$ can have in a graph with $\ell$ edges.

Look at $N(K_k, \ell)$.

$$(bln_b)^2 \leq (2\ell)^b$$

$$n_b \leq (2\ell)^{\frac{b}{2}}/bl!$$

$$N(K_k, \ell) \leq (2\ell)^{\frac{k}{2}}$$

The last equation is tight for a complete graph. But what is $N(T, \ell)$ in general?

Look at $T$, a 5-star (one node with edges to 5 nodes around it). For this $T$, $\ell$-star leads to at most $\ell^5$ embeddings and $\sqrt{2\ell}$ clique leads to $\sqrt{2\ell}$ embeddings. The number has to do with the structure of $T$.

#### 2.1.1 Fractional Independent Set

To understand $N(T, \ell)$ for an arbitrary graph $T$, we need to define two numbers associated with the graph $T$. The first is the fractional independent set number. A fractional independent set of $T$ is a function $\psi : V(T) \to [0,1]$ such that for every edge, $e = \{u,v\}, \psi(u) + \psi(v) \leq 1$. The size of the fractional independent set is $\alpha(\psi) = \sum_{v \in V} \psi(v)$. We write $\alpha^*(T)$ to denote the size of the biggest fractional independent set. Note that $\alpha^*(T)$ can be computed by a linear program, and the integer version of this program simply computes the size of the largest independent set.

This is a generalization of independent set. This continuous optimization is easier to solve than the discrete case.

The dual of this linear program measures a different quantity associated with $T$, namely the fractional cover number. Say that a mapping of the edges $\phi : E(G) \to [0,1]$ is a fractional cover if for every vertex $v$, $\sum_{e \in v} \phi(e) \geq 1$, where the sum is taken over all edges $e$ that contain $v$. The size of the fractional cover is $\gamma(\phi) = \sum_{e} \phi(e)$, and we denote by $\gamma^*(T)$ the size of the smallest fractional cover. Then the linear programming duality theorem proves that $\alpha^*(T) = \gamma^*(T)$.

If $T$ is a triangle, we have that $\alpha^*(T) = 3/2$, corresponding to the fractional independent set that weights every vertex with 1/2. Similarly, if $K$ is a $k$-clique, $\alpha^*(K) = k/2$. Indeed, the examples above are special cases of the following theorem, proved by Freidgut and Kahn (based on an earlier work of Alon).

**Theorem 3** ([1, 3]). If $T$ has $m$ edges, $(\ell/m)\alpha^*(T) \leq N(T, \ell) \leq (2\ell)^{\alpha^*(T)}$.

**Proof** First we prove the upper bound. Let $\sigma$ be a uniformly random embedding from $T \to G$, where $G$ is a fixed graph with $l$ edges. We shall use $\sigma$ to define a distribution on the edges of $T$ with high entropy. Let $\phi$ be the fractional cover of size $\alpha^*(T)$, and let $S$ be a random edge of $T$, such that for every edge $e$, $\Pr[S = e] = \phi(e)/\alpha^*(T)$. Namely, we use the distribution given by $\phi$ (after normalization). Now think of $\sigma$
as being specified by the values of \(\sigma(v)\) for all vertices \(v\) of \(T\). Then, since \(\phi\) is a fractional cover, we have that for every vertex \(v\), \(\Pr[v \in S] \geq \sum_{e \in \phi(v)} \phi(e)/\alpha^*(T) \geq 1/\alpha^*(T)\).

By Shearer’s Lemma, \(E_S[H(\sigma_S)] \geq H(\sigma)/\alpha^*(T)\). On the other hand, for each edge \(e\), \(\sigma_e\) is supported on edges of \(G\), so \(H(\sigma_e) \leq \log(2\ell)\). Thus \((2\ell)^{\alpha^*(T)} \geq N(T, \ell)\).

Next we prove the lower bound (modulo rounding arguments). Let us construct \(G\) for which there are many embeddings of \(T\) into \(G\). Let \(\psi\) be a fractional independent set that achieves \(\alpha^*(G)\). We obtain \(G\) by replacing every vertex in \(T\) with an independent set of \((\ell/m)^{\psi(v)}\) vertices, and connecting every vertex in the independent set for \(u\) to every vertex in the independent set for \(v\) if and only if \(\{u, v\}\) is an edge of \(T\). Every edge of \(T\) thus contributes \((\ell/m)^{\psi(u)}\psi(v) \leq \ell/m\) edges to \(G\), and so \(G\) has at most \(\ell\) edges. You can get a homomorphism from \(T\) to \(G\) by mapping any vertex \(v\) to a vertex in the independent set corresponding to \(v\), so there are at least \((\ell/m)^{\psi(v)} = (\ell/m)^{\alpha^*(T)}\) such homomorphisms. ■

### 2.2 Intersecting Families of Graphs

Suppose \(\mathcal{F}\) is a family of subsets of \([n]\). We say that \(\mathcal{F}\) is intersecting if for every \(A, B \in \mathcal{F}\), \(|A \cap B| > 0\).

One example of a large intersecting family is the family of sets that contain 1. This family has size \(2^n/2\), and this is as large as you can make such a family (because only one of \(A, A^c\) may belong to \(\mathcal{F}\)).

#### Claim 4. If \(\mathcal{F}\) is intersecting, then \(|\mathcal{F}| \leq 2^n/2\).

The proof is very simple: for every set \(A, \mathcal{F}\) can contain either \(A\) or its complement, but not both.

Next, let us call a family \(\mathcal{F}\) \(k\)-intersecting if for every \(A, B \in \mathcal{F}\), \(|A \cap B| \geq k\). An obvious example of such a family is the family of sets that all contain \(\{1, \ldots, k\}\), which has size \(2^n/2^k\). Can one do better?

Let \(\mathcal{F} = \{A \subseteq [n] : |A| \geq n/2 + k/2\}\). Then every two sets of \(\mathcal{F}\) intersect in at least \(k\) elements, but the size of \(\mathcal{F}\) is \(\sum_{i=n/2+k/2}^{n} \binom{n}{i} \geq (2^n/2)(1 - O(k/\sqrt{n}))\).

Next, let us try to place some structure on the intersections. Let \(\mathcal{G}\) be a family of graphs on the vertex set \([n]\). We say \(\mathcal{G}\) is intersecting if for any two graphs \(T, K \in \mathcal{G}\), \(T \cap K\) has an edge. Then as before, \(\mathcal{G}\) is of size at most \(2^{\binom{n}{2}/2}\), which can be achieved with the family of all graphs that contain a designated edge.

Things get interesting if we ask for the intersections to have some structure. Say that \(\mathcal{G}\) is \(\triangledown\)-intersecting if for every \(T, K \in \mathcal{G}\), \(T \cap K\) contains a triangle. The trivial example gives a family of size \(2^{\binom{n}{2}/8}\), but perhaps there is some clever way to get a \(\triangledown\)-intersecting family that has size close to \(2^{\binom{n}{2}/2}\), as in the examples above?

Chung, Frankl, Graham and Shearer showed that no such example exists:

#### Theorem 5 ([2]). If \(\mathcal{G}\) is \(\triangledown\)-intersecting, then \(|\mathcal{G}| \leq 2^{\binom{n}{2}/4}\).

**Proof** For any subset \(R \subseteq [n]\), let \(G_R\) be the graph consisting of two disconnected cliques, one on \(R\) and the other on the complement of \(R\). Write \(\lvert G_R \rvert\) for the number of edges in \(G_R\). Then observe that since for every \(T, K \in \mathcal{G}\), \(T \cap K\) contains a triangle, it must be the case that \(T \cap K \cap G_R\) contains an edge. Thus, the family of graphs \(\{T \cup G_R : T \in \mathcal{G}\}\) is intersecting, and so has size at most \(2^{\lvert G_R \rvert}/2\).

Let us define \(S\) to be a uniformly random graph \(G_R\) obtained by picking a random subset \(R\) of size \(n/2\). Observe that for any edge, by symmetry, the probability that the edge is included in \(G_R\) is \(|G_R|/\binom{n}{2}\).

Let \(G\) be a uniformly random graph from \(\mathcal{G}\). Consider what happens when we restrict \(G\) to the information about the edges in \(S\). By Shearer’s Lemma and the fact that \(G_S\) is supported on an intersecting family, \(\lvert G_R \rvert - 1 \geq \mathbb{E}_S[H(G_S)] \geq \frac{|G_R|}{\binom{n}{2}} \log |\mathcal{G}|\). Thus,
\[ \log |G| \leq \binom{n}{2} - \frac{\binom{n}{2}}{|G_R|} \]

\[ = \binom{n}{2} - \frac{\binom{n}{2}}{2^{\binom{n}{2}}} \]

\[ = \binom{n}{2} - \frac{n(n-1)}{2^{(n/2)(n/2-1)}} \]

\[ = \binom{n}{2} - \frac{n-1}{n/2-1} \]

\[ \leq \binom{n}{2} - 2 \]

---

**References**

