Statement of trickle-down, log-concavity for matroids

Last Time:

Some topics we covered last time were:

- Generating polynomials
- Log-concave distributions and polynomials
- Log-concavity implies 1-spectral independence at all links (where link = conditional measure)
- Statement of the "local-to-global" theorem [..., Alev-Lau '18, ...] which shows 1-spectral independence implies a spectral gap
- We saw the proof given a certain lemma.

Goals for Today:

The goal of this lecture is to continue to prove rapid mixing for the basis exchange walk.

In particular, we will:

1. Finish proving the Local-to-Global Theorem
2. Show most of the proof that the uniform distribution on the basis of a matroid is log-concave
   - We will prove this using "Oppenheim's Trickdown", a foundational result in the field, which we will prove next time.

1 Local-to-Global

Last time, we almost proved this result, except for not proving one lemma. Instead of just proving that lemma, we will first introduce some definitions for context and to prove the statement more generally.

First, we define more general down and up operators. Previously, we defined down up operators to go from sets of size k to $k-1$ and vice versa.

Definition 1. Higher Order Down Up Operators

We define the down operator that goes from sets of size k to sets of size l as

$$(\delta_S D_{k\to l})_T = \frac{1}{\binom{k}{l}} \mathbb{1}(T \subset S)$$

where we drop $k-l$ elements uniformly at random.

Then, using the same Bayes rule calculation as before, the up operator from a set of size l to one of size k is

$$(\delta_T U_{l\to k})_S \propto \mu(S) \mathbb{1}(T \subset S)$$

where we sample from the posterior on S where we generate the set T using the down operator: $S \overset{D_{k\to l}}{\to} T$. Altogether then, we define our transition matrix as $P_{k\leftrightarrow l} = D_{k\to l}U_{l\to k}$ for $k > l$.

These definitions will allow us to state the Local-to-Global argument for more general distributions, where last time we only covered 1-SI distributions. First, recall the definition of spectral independence.

Definition 2. μ is C-SI if

$$\mathrm{Cov}_\mu(\mathbb{1}_S) \preceq C\,\mathrm{diag}(\mathbb{E}[\mathbb{1}_S])$$

Recall that if μ is log-concave, then μ is 1-SI. That is, each link measure μ_S is 1-SI for all S.

Theorem 1. *Suppose μ is C-SI at all links. Then*

$$1 - \lambda_2(P_{k\leftrightarrow l}) \geq \prod_{i=0}^{l-1}\left(1 - \frac{C}{k-i}\right)$$

Note that provided $k - l$ is order 1, we can crudely approximate the RHS via $\frac{1}{k^C}$. Further, note that this generalization indicates that the statement is most powerful for $C = 1$ and less so for larger C, though still indicating polynomial time mixing.

Exercise 1. Show there exists 2-SI measures where $P_{k\leftrightarrow k-1}$ is not ergodic.

Before proving the theorem, we will state a fact and state and prove a key lemma.

Fact 1. $\mathbb{E}[f \mid s_1] = U_{1\to k}f$

Lemma 1. *C-SI implies* $\mathrm{Var}[U_{1\to k}f] \leq \frac{C}{k}\mathrm{Var}[f]$

Proof. (of lemma) ($\implies$)

Note that the inequality is equivalent to $\lambda_2(P_{k\leftrightarrow 1}) \leq \frac{C}{k}$, essentially by definition. To see this, note that with $\mathbb{E}[f] = 0$, the inequality is equivalent to $\langle f, f\rangle_\mu = \langle f, D_{k\to 1}U_{1\to k}f\rangle_\mu = \langle U_{1\to k}f, U_{1\to k}f\rangle_\mu \leq \frac{C}{k}\langle f, f\rangle_\mu$ which is the variational representation of the second largest eignevalue. In fact, this second inequality is also equivalent to $C - SI$. We will only prove the forward direction below, however.

So now, to show the lemma, we would like to show that $C - SI$ implies $\lambda_2(P_{k\leftrightarrow 1}) \leq \frac{C}{k}$.

First, note that second eigenvalue is invariant to the order of down and up operators: $\lambda_2(P_{k\leftrightarrow 1}) = \lambda_2(P_{1\leftrightarrow k})$, where $P_{1\leftrightarrow k}$ is a walk on sets of size 1, which is ideal because the transition matrix for this up-down walk can be easily written explicitly. Observe:

$$(P_{1\leftrightarrow k})_{ij} = \frac{1}{k}\mu(j \in S \mid i \in S) = \frac{1}{k}\frac{\mu(i \in S, j \in S)}{\mu(i \in S)}$$

Define the matrix of marginals $M_1 = \mathrm{diag}(\mu_1)$ where we define each marginal as $\mu_1(i) = \frac{1}{k}\mu(i \in S)$. So then

$$M_1 P_{1\leftrightarrow k} = \frac{1}{k^2}\mu(i \in S, j \in S)$$

and subtracting away the top eigendirection gives us

$$M_1 P_{1\leftrightarrow k} - \mu_1\mu_1^T = \frac{1}{k^2}\mathrm{Cov}(\mathbb{1}_S) \preceq \frac{C}{k}M_1$$

where in the inequality we have applied spectral independence.

Thus, taking $\mathbb{E}_{\mu_1}[f] = 0$ WLOG, we have recovered

$$\langle f, Pf\rangle_\mu = \langle U_{1\to k}f, U_{1\to k}f\rangle_\mu \leq \frac{C}{k}\langle f, f\rangle_\mu$$

the variational representation of the second eigenvalue of the up-down walk, which we know is equal to the second eigenvalue of the down-up walk. So we are done.

Proof. (of Theorem 1):

We proceed as last time. Define $S = \{s_1, \ldots, s_k\}$

By the law of total variance,

$$
\begin{aligned}
\mathrm{Var}[f] &= \mathbb{E}[\mathrm{Var}[f \mid s_1]] + \mathrm{Var}[\mathbb{E}[f \mid s_1]] && (1)\\
&= \mathbb{E}[\mathrm{Var}[f \mid s_1]] + \mathrm{Var}[U_{1\to k} f] \qquad \text{[fact 1]} && (2)\\
&\leq \mathbb{E}[\mathrm{Var}[f \mid s_1]] + \frac{C}{k}\mathrm{Var}[f] \qquad \text{[lemma 1]} && (3)\\
&\implies \left(1 - \frac{C}{k}\right)\mathrm{Var}[f] \leq \mathbb{E}[\mathrm{Var}[f \mid s_1]] && (4)
\end{aligned}
$$

Applying the same argument inducitvely to the right-hand side yields the desired result.

In summary: Log-concavity implies 1-SI implies $\Omega(1/k)$ sepctral gap.

2 Log-Concavity of Generating Polynomial for Matroids

Recall a set of bases $B \subseteq \binom{[n]}{k}$ are the bases of a matroid M if it satisfies an exchange property: $\forall S, T \subseteq B$, $\forall s \in S \setminus T$, $\exists t \in T \setminus S$, such that $(S \setminus \{s\}) \cup \{t\} \in B$. It may be useful to think of B as a very special set of a graph where B is the edge set and $[n]$ is the vertex set.

Exercise 2. Use this to prove the down-up walk is ergodic.

We now present two theorems. We will prove the first theorem via using the second (trickle-down) and checking a base case.

Theorem 2 (ALOV '24)**.** *Define the uniform measure $\mu = Unif(B)$. Then the generating polynomial*

$$g_\mu(\lambda) = \sum_S \mu(S) \prod_{i \in S} \lambda_i$$

is log-concave, i.e. $\log(g_\mu)$ *is concave on* $\mathbb{R}^n_{>0}$.

Theorem 3. *(trickle-down) Suppose that $\{\mu_i\}$ are C-SI for all i and further suppose that $P_{k\leftrightarrow 1}$ is ergodic. Then μ is C'-SI for*

$$C' = \frac{C(k-1)}{k-1-C}$$

Corollary 1. $C = 1 \implies C' = 1$ *so if you have 1-SI at a link, we get 1-SI at the link above ans so on. In this sense, we can say that $1 - SI$ "trickles-down".*

Proof. (of Theorem 2)

We will apply theorem 3 (trickle-down) to matroids to prove theorem 2. Most of what we will do is check the base case (k=2). This consists of verifying the generating polynomial is log-concave (i.e. the measure is 1-SI for all links). Then, trickle-down gives us that all links of μ are 1-SI. In fact, all tilts $\mu * \lambda$ are 1-SI (and so log-concave).

Thus, it suffices to check the base case and ensure it satisfies 1-SI. We will first prove the result without external fields. That is, if $\mu = Unif(B)$, where $B \subseteq \binom{[n]}{2}$, then μ is 1-SI. This is the main task of this lecture.

We will prove the following key lemma. Once we have that B is a complete multipartite graph, we are almost done because it is simple to show such graphs have adjacency matrix with small second eigenvalue.

Lemma 2. *B is a complete multipartite graph (when we ignore isolated vertices)*

Proof. (of lemma 2) The key lemma follows from some observations.

Observation 1: Suppose we have edges $\{i,j\}$ and $\{k,l\}$. By the exchange property if we drop j from $\{i,j\}$ then either $\{i,k\} \in B$ or $\{i,l\} \in B$. In particular, the graph distance $d(i,l) \leq 2$ for every vertex l because either l is a neighbor of i or by the exchange property l is a neighbor of a neighbor of i. Therefore, $G = ([n], B)$ has diameter at most 2.

So, we are working with a particularly special graph. To prove it is multipartite, we use the following lemma.

Lemma 3. *Suppose i and k are not neighbors. If $i \sim j$ then $j \sim k$.*

Proof. (of lemma 3) The Graph has diameter 2 so there exists some neighbor l such that $i \sim l$ and $k \sim l$. Consider $\{k,l\}$. Dropping k, by the exchange property we must be able to add i or j and have it be in B. Because we already know i and k are not neighbors, we know that $j \sim k$, as desired.

Using Lemma 3, it is easy to see that B is indeed a complete multipartite graph. Consider all the vertices k such that i and k are not neighbors and call them a part. Then all the vertices to which i connects also connect to all of the vertices in that part, by lemma 3. Then repeating this argument gives us the desired result: that is, we have a multipartite graph. That is, Lemma 3 implies Lemma 2.

The next lemma is an important result from spectral graph theory.

Lemma 4. *Let A be the adjaceny matrix. $\lambda_2(A) \leq 0$.*

Proof. There are many ways to prove this. For example:

Consider $\mathbb{1}\mathbb{1}^T - A$. Then this is a block diagonal matrix (with each block consisting of all ones) where each block corresponds to the different parts of the multipartite graph (because we know each part of the graph has no connections within). Say there are r blocks. This is a rank r matrix with eigenvalues $0, \ldots, 0, n_1, \ldots, n_r$. The eigenvector associated with n_1 is $[1_{n_1}, 0, \ldots, 0]$.In particular, $\mathbb{1}\mathbb{1}^T \succeq 0$ so $A \preceq \mathbb{1}\mathbb{1}^T$ so $\lambda_2(A) \leq 0$ by the variational characterization.

At this point, it is worth recalling what our goal is. We want to prove 1-SI is equivalent to $\lambda_2(P_{2\leftrightarrow 1}) \leq \frac{1}{2}$. It is unclear at this point how the eigenvectors of P relate to the eigenvectors of A. So we need the following (key) lemma.

Lemma 5. *Consider A symmetric and with non-negative entries. Then $\lambda_2(A) < 0$ if and only if*

$$(x^T A x)(v^T A v) \leq (X^T A v)^2$$

for all $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^n_{>0}$.

Proof. First prove the forward direction.

$$\begin{bmatrix} v^T \\ x^T \end{bmatrix} A \begin{bmatrix} vx \end{bmatrix} = \begin{bmatrix} v^T A v & v^T A x \\ x^T A v & x^T A x \end{bmatrix}$$

The determinant of this matrix M is $(x^T A x)(v^T A v) - (X^T A v)^2$. $v^T A v > 0$ implies that M has a positive eigenvalue. But $\lambda_2(A) < 0$ implies that M has to have a negative eigenvalue. So the determinant $\lambda_1(M)\lambda_2(M) < 0$, yielding the result.

Now we prove the reverse direction. Note that by the PSD inequality

$$A \preceq \frac{A v v^T A^T}{v^T A v}$$

which is rank one matrix so $\lambda_2(A) \leq 0$ and we are done.

Using Lemma 5, $\lambda_2(A) < 0$ is equivalent to $\lambda_2(DAD) < 0$ for any D diagonal matrix with positive entries.

Again, we would like to show $\lambda_2(P_{2\leftrightarrow 1}) \leq \frac{1}{2}$. Looking at the active random walk $D^{-1}A$, with $D = \text{diag}(deg(i))$, we would like $\lambda_2(D^{-1}A) \leq 0$ which is equivalent to $\lambda_2(D^{-1/2}AD^{-1/2}) \leq 0$ which is equivalent to $\lambda_2(A) \leq 0$ (because eigenvalues are invariant to appropriate change of bases) which we know.

Finally, we would like to extend the result to the case with external fields. g_μ is a quadratic polynomial. The Hessian

$$\nabla_\lambda^2 \log(g(\lambda_1, \ldots, \lambda_n)) = \frac{g(\lambda_1, \ldots, \lambda_n)\nabla_\lambda^2 g(r\lambda_1, \ldots, \lambda_n) - (\nabla g)(\nabla g)^T}{g(\lambda_1, \ldots, \lambda_n)^2}$$

Using $g(z) = \frac{1}{2}z^T Az$ for some A and $\nabla g = Az$ and $\nabla^2 g = A$, then $\nabla^2 \log g \mid_1 \leq 0$ so

$$A \leq \frac{Azz^T A^T}{z^T Az}$$

and recall this as an equivalent characterization that $\lambda_2(A) \leq 0$. Putting it together, we showed the above inequality for some z, which shows $\lambda_2(A) \leq 0$, which implies the inequality for all z.

In total, given trickle-down, we proved log-concavity of the generating polynomial, so $1 - \lambda_2(P_{k\to k-1}) \leq \Omega(1/k)$ for all bases of matroids. In particular for a a connected graph, spanning trees are bases for the "graphic matroid" so we proved rapid mixing for the down-up walk on spanning trees. This is much faster and more practical than the algorithm we saw previously.