6.896 Probability and Computation February 22, 2011

Lecture 6

Lecturer: Constantinos Daskalakis *Scribe: Brian Basham, Matthew Johnson*

NOTE: The content of these notes has not been formally reviewed by the lecturer. It is recommended that they are read critically.

1 Graph Coloring

Given a graph $G = (V, E)$ and a set of colors $\{1, 2, \ldots, q\}$, we wish to sample a random coloring uniformly from all possible proper colorings on the vertices of G. A coloring is legal if for every edge $e = (v_1, v_2) \in E$, v_1 and v_2 have distinct colors.

We define Δ to be the maximal degree of any vertex of G.

- $q \geq \Delta + 1 \implies \exists$ a coloring of G. We can see this by simply choosing a color for each vertex in some fixed order, and at every step there is at least one color not among the neighbors of v.
- $q < \Delta$: In general, it is an NP-Hard problem to decide if there is a coloring of G or not.
- $q = \Delta$: Brook's Theorem states that there exists a coloring of G iff there is no $\Delta + 1$-clique in G and $\Delta > 2$, or if there is no odd cycle in G and $\Delta = 2$.

From this point on we will focus on the case $q \geq \Delta + 1$, so we know that there exists at least one coloring of G. We now wish to sample a coloring uniformly. We start by considering the natural Markov chain on the colorings of a graph G.

Definition 1. *The* natural Markov chain *on a set of colorings of $G = (V, E)$ and a set of colors $\{1, 2, \ldots, q\}$ is defined as follows. We let X_0 be an arbitrary coloring of G. X_{t+1} is obtained by X_t by choosing $v \in V$ uniformly at random and choosing a uniform $c \in \{1, 2, \ldots, q\}$, changing the color of v to c if the result would be a legal coloring and otherwise keeping the colors of all vertices the same.*

This Markov chain is symmetric because each possible change is reversible with the same probability, and is aperiodic because it contains self-loops. However, it is not irreducible. Consider $\Delta = 2, q = 3$, and G is a complete graph on 3 vertices. Any coloring cannot be transformed to another coloring using this process.

Exercise (1 point): Prove that the natural Markov chain is irreducible if $q \geq \Delta + 2$.

Conjectures (10 points each):

(a) If $q \geq \Delta + 2$ the natural Markov chain has $\tau_{mix} = O(n \log n)$

(b) If $q = \Delta + 1, \exists$ a Markov chain with polynomial mixing time.

Today: The natural Markov chain has $\tau_{mix} = O(n \log n)$ for $q \geq 4\Delta + 1$.
Tomorrow: We will improve the bound to $q \geq 2\Delta + 1$ using Path Coupling.

Theorem 1. *The mixing time of the natural Markov chain is $\tau_{mix} = O(n \log n)$ for $q \geq 4\Delta + 1$*

Proof:
We use coupling. We choose arbitrary colorings X_0 and Y_0 of G, and we couple (X_t, Y_t) by picking the same v and t uniformly at random at all times t.

We denote by $d_t = d(X_t, Y_t)$ the number of vertices where X_t and Y_t differ in color.

There are three types of possible moves. Good moves (those that decrease d_t), bad moves (that increase d_t), and neutral moves (d_t doesn't change).

- Good Moves ($d_{t+1} = d_t - 1$): This occurs when v is a vertex that disagrees and c is a color that does not appear in the neighborhood of v in either X_t or Y_t. There are at most 2Δ colors in the neighborhood of v in either coloring, so there are at least $d_t(q - 2\Delta)$ good pairs of v, c.
- Bad Moves ($d_{t+1} = d_t + 1$) This occurs when v is a vertex that is the same color in X_t and Y_t and c is a color that appears among the neighbors of v in only one of X_t or Y_t but not both. We bound this by counting the disagreeing neighbors. If c appears among the neighbors of v in only one coloring, then that vertex must disagree in X_t and Y_t. There are at most d_t disagreeing vertices each with at most Δ neighbors and two choices of colors, for a total of at most $2\Delta d_t$ bad moves.
- Neutral moves ($d_{t+1} = d_t$) Everything that isn't a Good or Bad move is a Neutral move.

We notice that (#Good Moves − #Bad Moves $\geq q - 4\Delta$)

$$\begin{aligned}
E[d_{t+1}|(X_t, Y_t)] &= d_t - \frac{\#\text{Good Moves}}{qn} + \frac{\#\text{Bad Moves}}{qn} \\
&\leq d_t\left(1 - \frac{q - 4\Delta}{qn}\right) \\
\implies E[d_t|(X_0, Y_0)] &\leq d_0\left(1 - \frac{q - 4\Delta}{qn}\right)^t \\
&\leq n e^{-t\frac{q-4\Delta}{qn}} \\
&\leq \epsilon \text{ when } t \geq \frac{q}{q - 4\Delta} n\left(\log n + \log \frac{1}{e}\right)
\end{aligned}$$

By Markov's Inequality,

$$Pr[X_t \neq Y_t|(X_0, Y_0)] = Pr[d_t \geq 1|(X_0, Y_0)] \leq E[d_t|(X_0, Y_0)] \leq \epsilon$$

for $t \geq \frac{q}{q-4\Delta} n(\log n + \log \frac{1}{e})$

So $\tau(\epsilon) = \frac{q}{q-4\Delta} n(\log n + \log \frac{1}{e})$,

$\tau_{mix} = O(\frac{q}{q-4\Delta} n \log n) = O(n \log n)$ for $q \geq 4\Delta + 1$. □

2 Path Coupling

We will develop the method of Path Coupling, which we will use to achieve a mixing time of $O(n \log n)$ when $q \geq 2\Delta + 1$.

Definition 2. *: A* pre-metric *on Ω is a connected, undirected graph with positive edge weights such that all edges are shortest paths. More specifically, if $x, y \in \Omega$ are adjacent in the pre-metric, the weight of the edge between them $w(x, y)$ should be equal to the least weight of any path from x to y in the pre-metric.*

We can extend any pre-metric to a metric on Ω by considering the shortest path distances among edges of the pre-metric. We let the induced metric be d.

Idea of Path Coupling: We define a coupling only for pairs of states that are adjacent in the pre-metric such that the distance under our new metric is expected to reduce by some constant factor. We then extend this to a full coupling that still reduces the expected distance by the same factor.

Theorem 2. *Path Coupling Theorem*

Suppose $\exists$ *coupling* $Pr[X', Y'|X, Y]$, *defined only for pairs of states* (X, Y) *that are adjacent in the pre-metric, such that*

$$E[d(X', Y')|X, Y] \leq (1 - \alpha)d(X, Y) \qquad (*)$$

Then we can extend this coupling to a full coupling such that $(*)$ *holds for all* $(X, Y) \in \Omega$.

Proof: Let x and y be arbitrary states that are non-adjacent in the pre-metric. We fix an arbitrary shortest path between these states,

$$x = z_0 \to z_1 \to z_2 \to \cdots \to z_k = y$$

We need to specify $Pr[Z'_0, Z'_k|z_0, z_k]$ that is a valid coupling. We first sample Z'_0, Z'_1 from $Pr[Z'_0, Z'_1|z_0, z_1]$. We then sample a value of Z'_2 from the marginal distribution of $Pr[Z'_1, Z'_2|z_1, z_2]$ conditioning on the sampled value of Z'_1. We then sample a value for $Z'_3, Z'_4, \ldots Z'_k$ conditioning on the previous value each time, and we have now sampled $(Z'_0, Z'_K) = (X', Y')$.

Claim 1. *The above procedure is i) well defined and ii) a valid coupling.*

Proof: We show by induction that $Pr[Z'_i = w] = P(z_i, w) \forall i, w \in \Omega$, where $P(z_i, w)$ is the Markov Chain transition probability from z_i to w. This means that $z_i \to Z'_i$ is a valid Markov Chain step.

For Z'_0, Z'_1 this is true because $Pr[Z'_0, Z'_1|z_0, z_1]$ defines a valid coupling.

We now assume the claim is true for Z'_i, and we will show that it is true for Z'_{i+1}.

$$\begin{aligned}
Pr[Z'_{i+1} = w] &= \sum_{z'_i} P(Z'_i = z_i) \frac{Pr[Z'_i = z'_i, Z'_{i+1} = w|z_i, z_{i+1}]}{\sum_{w^*} Pr[Z'_i = z'_i, Z'_{i+1} = w^*|z_i, z_{i+1}]} \\
&= \sum_{z'_i} P(z_i, z'_i) \frac{Pr[Z'_i = z'_i, Z'_{i+1} = w|z_i, z_{i+1}]}{\sum_{w^*} Pr[Z'_i = z'_i, Z'_{i+1} = w^*|z_i, z_{i+1}]} \text{ by induction.} \\
&= \sum_{z'_i} P(z_i, z'_i) \frac{Pr[Z'_i = z'_i, Z'_{i+1} = w|z_i, z_{i+1}]}{P(z_i, z'_i)} \text{ because } Pr[Z'_i, Z'_{i+1}|z_i, z_{i+1}] \text{ is a valid coupling.} \\
&= \sum_{z'_i} Pr[Z'_i = z'_i, Z'_{i+1} = w|z_i, z_{i+1}] \\
&= P(z_{i+1}, w) \text{ because } Pr[Z'_i, Z'_{i+1}|z_i, z_{i+1}] \text{ is a valid coupling.}
\end{aligned}$$

□

Thus *ii*) is established, and we have implicitly shown *i*). We also notice that the distribution of every (Z'_i, Z'_{i+1}) is identical to $Pr[Z'_i, Z'_{i+1}|z_i, z_{i+1}] \forall i$. Now we can finish the proof.

$$\begin{aligned}
E[d(Z'_0, Z'_k)|z_0, z_k] &\leq E[\sum_{i=0}^{k-1} E[d(Z'_i, Z'_{i+1})|z_0, z_k] \\
&\leq \sum_{i=0}^{k-1} E[d(Z'_i, Z'_{i+1}|z_0, z_k] \text{ by linearity of expectation.} \\
&\leq \sum_{i=0}^{k-1} (1 - \alpha)d(z_i, z_{i+1}) \\
&= (1 - \alpha) \sum_{i=0}^{k-1} d(z_i, z_i + 1) \\
&= (1 - \alpha)d(z_0, z_k)
\end{aligned}$$

as desired. We have established the Path Coupling Theorem. □