Introduction

Today we continue to show applications of the coupling technique. We are dealing with the problem of generating k-colorings.

1 k-colorings

Problem: k-coloring

Input: Graph G, integer k

Output: Random k-coloring (from the uniform distribution)

By a k-coloring we mean a function $C : V \to \{0, \ldots k-1\}$ such that if $(u, v) \in E(G)$, $C(u) \neq C(v)$.

Now we describe algorithms for general graphs. Later we will consider 3-colorings of the Cartesian lattice.

We assume a k-coloring is given. (We need this assumption since all the algorithms start with some k-coloring; in general finding if a k-coloring exists is NP-complete.)

The first Markov chain on the space of k-colorings of G is SPR (Single Point Recoloring). The transition rule is:

Repeat:

1. Choose uniformly $(p, c) \in V(G) \times \{0, \ldots k-1\}$.
2. Recolor p with color c if possible

For most interesting graphs, you can prove the state space is connected. However this is not true in general. Note that this Markov chain is symmetric: we can move from one coloring to another only if they differ at exactly one vertex, and the probability to move from one to another is $1/kn$ in this case.

A second Markov chain on the space of k-colorings of G is the Swendsen-Wang Kotesky (SWK). The transition rule is:

Repeat:

1. Choose uniformly $(c_1, c_2) \in \{0, \ldots k-1\} \times \{0, \ldots k-1\}$.
2. For each $c_1 - c_2$ component (i.e., component of the graph induced in G by the vertices colored c_1 or c_2) independently decide whether to recolor (switch the colors)

A third Markov chain is a SWK variant:

Repeat:

1. Uniformly choose $(p, c) \in V(G) \times \{0, \ldots k-1\}$.
2. If p is colored c', flip the $c - c'$ component containing p.

Theorem : (Jerrum) If $k \geq 2\Delta$, SPR is rapidly mixing, where Δ is the maximum degree of G.
Proof :

By coupling. Here the coupling rule (which we will define soon) will depend on the current configuration. Let MC^1 be the original Markov chain and MC^2 the coupled one.

At time t we will have two coloring: C_t^1 and C_t^2. Let $dist(C^1, C^2)$ be the number of vertices colored differently. So $0 \leq dist \leq |V| = n$. By Δ^t $dist(C^1, C^2)$ we denote $dist(C_{t+1}^1, C_{t+1}^2) - dist(C_t^1, C_t^2)$.

Our goal is to prove for the coupling rule which we will define soon that for all t,

$$E[\Delta\ dist(C^1, C^2)] \leq 0$$

and assuming $C^1 \neq C^2$, for all t,

$$P[\Delta\ dist(C^1, C^2) \neq 0] \geq \frac{1}{n}$$

Assuming we proved these two bounds, it will follow from a general martingale argument (which will appear later in the class) that the expected number of steps to get to $dist(C^1, C^2) = 0$ is $O(n^4)$. For intuition only, note that $dist$ does a random walk on a path graph of length n, moves with positive probability, and in expectation is more likely to decrease than increase.

Before defining the coupled random walk, we need some notation. Let $A = \{p \in V(G) | C^1(p) = C^2(p)\}$ (the set of vertices whose colors agree in the two colorings). Let $D = V(G) \setminus A$ (the set of vertices whose colored differently in the two colorings). By A_p we denote $A \cap N(p)$ (where as usual $N(p)$ is the set of neighbors of p in G) and by D_p, $D \cap N(p)$.

To define the coupled Markov chain MC^2, we do as follows: first we uniformly choose $(p, c) \in V(G) \times \{0, \ldots k-1\}$. If $p \in D$, use the same (p, c) to do the single point recoloring in MC^2.

If $p \in A$, the rule is more complicated. Let $\{c_1, \ldots c_u\}$ be the set of colors s such that in C^1 the vertex p has a neighbor colored s, but in C^2, no neighbor of p is colored s. Similarly, let $\{q_1, \ldots q_r\}$ be the set of colors s such that in C^2 the vertex p has a neighbor colored s, but in C^1, no neighbors of p is colored s. Note that since the vertices in A_p have the same color in both C^1 and C^2, $u \leq |D_p|$ and $r \leq |D_p|$.

Let $j = \min(u, r)$. If $c = c_i$ for $1 \leq i \leq j$, use (p, q_i) for the transition in MC^2. Similarly, if $c = q_i$ for $1 \leq i \leq j$, use (p, c_i) for the transition in MC^2. For all other c, use the same (p, c) for the transition in MC^2 as in MC^1.

A seed (move) in $V(G) \times \{0, \ldots k-1\}$ is *good* if it decreases $dist(C^1, C^2)$ and *bad* if it increases $dist(C^1, C^2)$. As all seeds are equally likely, to prove that $E[\Delta\ dist(C^1, C^2)] \leq 0]$ it is enough to show that the number of good seeds is at least the number of bad seeds.

Let $p \in D$ and let $c \in \{0, \ldots k-1\}$ be a color such that no neighbor of p is colored c in any of C^1 and C^2. Then the seed (p, c) is good. The number of such colors c is at least $k - (2d(p) - |A_p|)$, where $d(p)$ is the degree of p. We conclude that the total number of good seeds is at least

$$\sum_{p \in D} k - (2d(p) - |A_p|) \geq \sum_{p \in D} |A_p|$$

as $k \geq 2\Delta \geq 2d(p)$.

Note that if $p \in D$ no seed (p, c) can be bad. Let now $p \in A$. We are going to argue that the number of bad seeds having p as the vertex is at most by D_p.

First note that if $c \notin \{c_1, \ldots c_u\} \cup \{q_1, \ldots q_r\}$, the seed (p, c) is not bad: in both Markov chains the same thing (p gets recolored to c or stays the same color) happens. Also (p, c_i) is not bad if

$1 \leq i \leq j$. Indeed, for $1 \leq i \leq j$, in MC^1 no recoloring takes place, and since we use the (p, q_i) for the transition in MC^2, no recoloring takes place in MC^2 either. So the number of bad seeds is at most $u + r - j = \max(u, r) \leq |D_p|$. In conclusion the total number of bad seeds is at most:

$$\sum_{p \in A} |D_p| = \sum_{p \in D} |A_p|,$$

since we are counting the number of edges of G with one endpoint in A and one endpoint in D. So we have proved that $E[\Delta\ dist(C^1, C^2)] \leq 0]$.

Note that if $C^1 \neq C^2$, there is at least one good seed and therefore $P[\Delta\ dist(C^1, C^2) \neq 0] \geq \frac{1}{kn} \geq \frac{1}{2n^2}$. This gives ust that the mixing time is polynomial.

□

We have time to mention the important problem of generating random colorings of the two dimensional Cartesian lattice. For $k \geq 8$ the theorem above shows that SPR is rapidly mixing, while for $k = 3$ a variant of SPR works too. The problem is open for k=4, 5, 6, and 7.