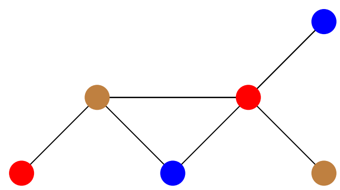
Markov Chains and Algorithmic Applications: WEEK 11

1 Application: Graph coloring

Let $G = (V, E)$ be a graph with vertex set V ($|V| = N$) and edge set E. We want to color each vertex of the graph with one of the q colors at our disposal such that a vertex's color differs from that of all its neighbors, as seen below:



More formally, let S be the set of all possible color configurations on G and $x = (x_v,\ v \in V) \in S$ a particular color configuration. A *proper q-coloring* of G is any configuration x such that $\forall v, w \in V$, if $(v, w) \in E$ then $x_v \neq x_w$.

Our aim: to sample uniformly amongst the proper q-colorings of G. In other words, we want to sample from the distribution

$$\pi(x) = \frac{\mathbb{1}_{\{x \text{ is a proper } q\text{-coloring}\}}}{Z = \sharp \text{ proper } q\text{-colorings}}, \quad x \in S$$

Remark 1.1. Let $\Delta = \max_{v \in V} \deg(v)$. If $q \geq \Delta + 1$, then there exists at least one proper q-coloring.

In what follows, we are going to restrict our analysis to graphs satisfying $q > 3\Delta$.

One way to sample from π is by using the following algorithm:

1. Start from a proper q-coloring $x \in S$.
2. Select a vertex $v \in V$ uniformly at random.
3. Select a color $c \in \{1, \dots, q\}$ uniformly at random.
4. If c is an allowed color at v, then recolor v (i.e. set $x_v = c$); do nothing otherwise.
5. Repeat steps 2, 3 and 4.

Remark 1.2. Since the algorithm started from a proper q-coloring $x \in S$, every visited state is also a proper q-coloring.

Remark 1.3. The algorithm could also be used to *find* a proper q-coloring on G. Indeed, if we start the algorithm in a state $x' \in S$ that is not a proper coloring, the algorithm ensures that eventually a proper coloring will be reached.

Definition 1.4. Let $x, y \in S$ be two color configurations. We write $x \sim y$ if x and y differ in at most one vertex.

Remark 1.5. The algorithm is actually an instance of the Metropolis-Hastings algorithm:

1. Let the base chain be $\psi_{xy} = \begin{cases} \frac{1}{Nq} & y \sim x, y \neq x, \\ \frac{1}{q} & y = x, \\ 0 & \text{otherwise.} \end{cases}$

   ψ is aperiodic (due to self-loops) and satisfies $\psi_{xy} > 0$ iff $\psi_{yx} > 0$ (due to symmetry). Moreover, the condition $q > 3\Delta$ ensures that ψ is irreducible.

2. $a_{xy} = \min\left(1, \frac{\pi_y}{\pi_x}\right) = \min\left(1, \frac{\mathbb{1}_{\{y \text{ is a proper } q\text{-coloring}\}}/Z}{1/Z}\right) = \mathbb{1}_{\{y \text{ is a proper } q\text{-coloring}\}}$.
   (NB: we already know that x is a proper q-coloring).

3.

$$p_{xy} = \begin{cases} \psi_{xy} a_{xy} & y \neq x, \\ 1 - \sum_{z \in S \setminus x} \psi_{xz} a_{xz} & y = x \end{cases}$$
$$= \begin{cases} \frac{1}{Nq} \mathbb{1}_{\{y \text{ is a proper } q\text{-coloring}\}} & y \sim x,\ y \neq x, \\ 1 - \frac{1}{Nq} \sharp\{z \sim x,\ z \neq x,\ z \text{ proper } q\text{-coloring}\} & y = x, \\ 0 & \text{otherwise.} \end{cases}$$

1.1 Convergence rate analysis

The *mixing time* of this chain is $T_\epsilon = \inf\{n \geq 1 : \max_{x \text{ proper } q\text{-coloring}} \|P_x^n - \pi\|_{\text{TV}} \leq \epsilon\}$.

Theorem 1.6. If $q > 3\Delta$, then for all proper q-colorings x, $\|P_x^n - \pi\|_{\text{TV}} \leq N e^{-\frac{n}{N}\left(1 - \frac{3\Delta}{q}\right)}$, implying that

$$T_\epsilon \leq \frac{1}{1 - \frac{3\Delta}{q}} N \left(\log(N) + \log\left(\frac{1}{\epsilon}\right)\right)$$

Proof. Let $(X_n, n \geq 0)$ be a Markov chain on S starting at $X_0 = x$ (a proper q-coloring) and evolving according to P. Let $(Y_n, n \geq 0)$ be a Markov chain on S starting at $Y_0 \sim \pi$ and also evolving according to P.

We will couple X and Y as follows:

1. Select a vertex $v \in V$ uniformly at random.
2. Select a color $c \in \{1, \ldots, q\}$ uniformly at random.
3. Update X at vertex v if c is an allowed color.
   Update Y at vertex v if c is an allowed color.

Definition 1.7. The *Hamming distance* between two colorings x and y is the number of positions in which x and y disagree:

$$d(x, y) = \sum_{v \in V} \mathbb{1}_{\{x_v \neq y_v\}}$$

By a coupling argument seen in previous lectures, we have

$$\|P_x^n - \pi\|_{\text{TV}} \leq \mathbb{P}(X_n \neq Y_n) = \mathbb{P}(d(X_n, Y_n) \geq 1) \leq \mathbb{E}(d(X_n, Y_n)),$$

where the last inequality is obtained by using the Markov inequality.

All that is left to do now is to upper bound $\mathbb{E}(d(X_n, Y_n))$. We will do so using two inductions:

1. Assume first that $d(X_0, Y_0) = 1$, i.e. X_0 and Y_0 differ at one vertex only, and let v be that vertex. Due to the coupling, at most one vertex can change color per transition, hence $d(X_1, Y_1) \in \{0, 1, 2\}$ and

$$\begin{aligned} \mathbb{E}(d(X_1, Y_1)) &= 0 \cdot \mathbb{P}(d(X_1, Y_1) = 0) + 1 \cdot \mathbb{P}(d(X_1, Y_1) = 1) + 2 \cdot \mathbb{P}(d(X_1, Y_1) = 2) \\ &= (1 - \mathbb{P}(d(X_1, Y_1) = 0)) + \mathbb{P}(d(X_1, Y_1) = 2) \end{aligned}$$

$d(X_1, Y_1) = 0$ if and only if vertex v is chosen (with probability $\frac{1}{N}$) and that the color c chosen is allowed in both chains X and Y, hence

$$\mathbb{P}(d(X_1, Y_1) = 0) = \frac{1}{N} \cdot \frac{\sharp \text{ allowed colors at } v}{q} \geq \frac{1}{N} \cdot \frac{q - \Delta}{q}$$

$d(X_1, Y_1) = 2$ if and only if the vertex w chosen is a neighbor of v (because $p_{ij} = 0$ when $i \not\sim j$) and that either X or Y is recolored (but not both). The latter only happens when the chosen color c satisfies $c = x_v$ or $c = y_v$, so we have

$$\mathbb{P}(d(X_1, Y_1) = 2) \leq \frac{\Delta}{N} \cdot \frac{2}{q}$$

Gathering both estimates together, we obtain

$$\mathbb{E}(d(X_1, Y_1)) \leq \left(1 - \frac{1}{N}\frac{q - \Delta}{q}\right) + \frac{\Delta}{N}\frac{2}{q} = 1 - \frac{1}{N}\left(1 - \frac{3\Delta}{q}\right)$$

2. Suppose now that $d(X_0, Y_0) = r$. Since P describes an irreducible Markov chain, there exists a sequence of $r - 1$ states $Z_0^{(1)}, \ldots, Z_0^{(r-1)}$ such that

$$p_{X_0 Z_0^{(1)}} p_{Z_0^{(1)} Z_0^{(2)}} \cdots p_{Z_0^{(r-1)} Y_0} > 0,$$

$$d\left(X_0, Z_0^{(1)}\right) = d\left(Z_0^{(1)}, Z_0^{(2)}\right) = \cdots = d\left(Z_0^{(r-1)}, Y_0\right) = 1$$

This implies that

$$\begin{aligned}
\mathbb{E}(d(X_1, Y_1)) &\leq \mathbb{E}\left(d\left(X_1, Z_1^{(1)}\right)\right) + \mathbb{E}\left(d\left(Z_1^{(1)}, Z_1^{(2)}\right)\right) + \cdots + \mathbb{E}\left(d\left(Z_1^{(r-1)}, Y_1\right)\right) \\
&= r\left(1 - \frac{1}{N}\left(1 - \frac{3\Delta}{q}\right)\right)
\end{aligned}$$

3. This inequality is valid between times 0 and 1, but it also holds between times n and $n + 1$:

$$\mathbb{E}(d(X_{n+1}, Y_{n+1}) \,|\, d(X_n, Y_n) = r) \leq r\left(1 - \frac{1}{N}\left(1 - \frac{3\Delta}{q}\right)\right)$$

From the above, we deduce that

$$\begin{aligned}
\mathbb{E}(d(X_{n+1}, Y_{n+1})) &\leq \left(1 - \frac{1}{N}\left(1 - \frac{3\Delta}{q}\right)\right)\mathbb{E}(d(X_n, Y_n)) \\
\Longrightarrow \mathbb{E}(d(X_n, Y_n)) &\leq \mathbb{E}(d(X_0, Y_0))\left(1 - \frac{1}{N}\left(1 - \frac{3\Delta}{q}\right)\right)^n \\
&\leq N e^{-\frac{n}{N}\left(1 - \frac{3\Delta}{q}\right)}
\end{aligned}$$

which completes the proof. □