Path Coupling

Seminar in Random Walks

Omer Paneth

November 16, 2009

- Uniform sampling using the Gibbs sampler
- Fast convergence using a coupling argument
- No general coupling for Gibbs sampler
- for q-coloring we could show fast convergence when $q \geq 2\Delta^2$

- Intuition for path coupling
- Path coupling lemma
- A random q-coloring for $q = 2\Delta + 1$
- Path coupling for Gibbs sampler
- Linear extensions of a partial order

- **Path Coupling**
  General method for defining a good coupling
- **Idea**
  Only couple states that are already close
- **Example:** Random q-coloring

- M - Markov chain on a set Ω
- $D(\Omega, R)$ - Weighted connected undirected graph
- $\delta(x, y)$ - Weight of the shortest path $x \leftrightsquigarrow y$ in D

Define a stochastic process C on $(X_t, Y_t) \in R \subset \Omega \times \Omega$ s.t.:

- Marginal distributions of X_t, Y_t are copies of M
- $\exists \beta < 1, \forall (x, y) \in R$,

$$\mathbb{E}[\delta(X_1, Y_1) | (X_0, Y_0) = (x, y)] \leq \beta \cdot \delta(x, y)$$

Path Coupling Lemma

Extend C to a coupling on $\Omega \times \Omega$:
For $(x, y) \in \Omega \times \Omega$, let $x = x_1, x_2, \ldots, x_k = y$ be the shortest path $x \leftrightsquigarrow y$ in D, i.e., $\forall i (x_i, x_{i+1}) \in R$

- First pick x'_1 according to M given x_1
- For i in $1..k-1$:

  Pick x'_{i+1} according to the marginal distribution $C(x_i, x_{i+1})$ given x'_i
- Move to state $(x'_1, x'_k) = (x', y')$

Lemma. $\forall (x, y) \in \Omega \times \Omega$, $\mathbb{E}[\delta(x', y')] \leq \beta \cdot \delta(x, y)$

Proof.

$$\begin{aligned}
\mathbb{E}[\delta(x', y')] &\leq \mathbb{E}\left[\sum_{i=1}^{k-1} \delta(x'_i, x'_{i+1})\right] && \text{(minimiality of } \delta\text{)} \\
&= \sum_{i=1}^{k-1} \mathbb{E}\left[\delta(x'_i, x'_{i+1})\right] && \text{(linearity of exp.)} \\
&\leq \sum_{i=1}^{k-1} \beta \cdot \delta(x_i, x_{i+1}) && \text{(definition of } \beta\text{)} \\
&= \beta \cdot \delta(x, y) && \text{(definition of } \delta\text{)}
\end{aligned}$$

□

Random q-coloring

For a graph $G(V,E)$ of maximal degree Δ and a set C of q colors, define a chain on C^V:

- Pick at random $v \leftarrow V$, $c \leftarrow C$
- If v can be colored in c color it
  Else do nothing

This is not exactly a Gibbs sampler but it still converge to the uniforms distribution

- We will use Hamming distance - $H(X,Y)$
- Consider the graph $D(\Omega, R)$, where $\Omega = C^V$, $R = \{(X,Y) \in \Omega \times \Omega | H(X,Y) = 1\}$

Coupling for $(X,Y) \in R$:

- Pick at random $v \leftarrow V$, $c \leftarrow C$
- Try to move according to v, c in both X and Y

Assuming X, Y differ only in the color of v_0

- If v is v_0 then $H(X,Y)$ can decrease to 0
  $\Pr[v = v_0 \text{ and } c \text{ is a legal color}] \geq \frac{1}{n}(1 - \frac{\Delta}{q})$
- If v is adjacent to v_0 then $H(X,Y)$ can increase to 2
  $\Pr[(v, v_0) \in E \text{ and } c \in \{X(v), Y(v)\} \leq \frac{\Delta}{n}\frac{2}{q}$
- Else, $H(X,Y)$ stays 1, so:

$$\mathbb{E}[\delta(X', Y')] < 1 - \frac{1}{n}\left(1 - \frac{\Delta}{q}\right) + \frac{\Delta}{n}\frac{2}{q} = 1 - \frac{q - 3\Delta}{n \cdot q}$$

- For $q \geq 3\Delta + 1$ we can set β to $1 - \frac{1}{n \cdot q}$
- D has diameter n
- By path coupling $\mathbb{E}[H(X_t, Y_t)] \leq \beta^t \cdot n$
- $\Pr[X_t \neq Y_t] \leq \mathbb{E}[H(X_t, Y_t)] \leq \beta^t \cdot n$
- By coupling lemma:

$$\tau(\epsilon) \leq \left\lceil \frac{\ln(n \cdot \epsilon^{-1})}{\ln \beta^{-1}} \right\rceil = \left\lceil \frac{\ln(n \cdot \epsilon^{-1})}{\ln\left(1 + \frac{1}{n \cdot q - 1}\right)} \right\rceil$$

$$\tau(\epsilon) = O(n^2 \log(n \cdot \epsilon^{-1}))$$

Random q-coloring - Coupling 2

We can improve the bound on q by defining a better coupling:

- Pick at random $v \leftarrow V$, $c \leftarrow C$
- If $(v, v_0) \in E$ and $c \in \{X(v_0), Y(v_0)\}$
  - In X, try coloring v with rand color in $\{X(v_0), Y(v_0)\}$
  - In Y, try coloring v with the other color
- Else, color $X(v), Y(v)$ with c

Notice that the process is still a coupling, since the probability to color v with color $X(v_0)$ has not changed (same for $Y(v_0)$)

Random q-coloring - Analysis 2

- $\Pr[(v, v_0) \in E$ and $X'(v) \neq Y'(v)] \leq \frac{1}{2}\frac{\Delta}{n}\frac{2}{q}$
- Similar to the previous case we get:

$$\begin{aligned} \mathbb{E}[\delta(X', Y')] &< 1 - \frac{1}{n}\left(1 - \frac{\Delta}{q}\right) + \frac{1}{2}\frac{\Delta}{n}\frac{2}{q} \\ &= 1 - \frac{q - 2\Delta}{n \cdot q} \end{aligned}$$

And the same result holds for $q \geq 2\Delta + 1$

Path Coupling for Gibbs Sampler

For states X, Y where $H(X, Y) = 1$ define the following coupling:

- Pick at random $v \leftarrow V$
- Find $\nu_{X,v}$ - the distributions of possible colors for v in X
- Similarly fine $\nu_{Y,v}$
- Color v in X and Y according to the optimal coupling between $\nu_{X,v}$ and $\nu_{Y,v}$

It is left to find out if we can find $\beta < 1$ as required

Linear Extensions of a Partial Order

Let P be a partial order on $[n]$. A linear extension of P is a total order (permutation) which is consistent with P.

The problem: Given a partial order P pick a random linear extension of P.

To sample uniformly from all linear extensions of P we will use the following chain on any linear extension X of P:

- Pick at random $c \leftarrow \{0, 1\}$
- Pick p according to some distribution f on $[n-1]$
- If $c = 1$ and transposing $X(p)$ and $X(p+1)$ is consistent with P, transpose them
  Else, do nothing

The chain defined is interesting because:

- It is irreducible, since we can get from every legal permutation X to Y by ”moving” every element of X to it’s new place, according to the order of Y
- It is aperiodic, since is contains self loops
- The transition matrix is symmetric so the uniform distribution is reversible and stationery

It remains to show that the chain is rapidly mixing

Linear Extensions - Notation

Denote:

- Ω is the set of all linear extensions of P
- $\sigma(i,j)$ for $0 < i < j \leq n$ is the transposition operation
- For $\sigma(i,j)$, $j - i$ is the transposition width
- $D(\Omega, R)$, where $R = \{(X,Y) | \exists i < j, \sigma(i,j)X = Y\}$
- The weight of an edge in R is it's transposition width
- $\delta(X,Y)$ is the weight of the shortest path $X \leftrightsquigarrow Y$ in D

For all $X, Y \in R$ s.t. $\sigma(i,j)X = Y$ define the following coupling:

- Pick random $c_0 \in \{0,1\}$
- Pick p according to f
- If $p = i$ and $p+1 = j$then $c_1 = 1 - c_0$ else $c_1 = c_0$
- If $c_0 = 1$ try to apply $\sigma(p, p+1)$ on X
- If $c_1 = 1$ try to apply $\sigma(p, p+1)$ on Y

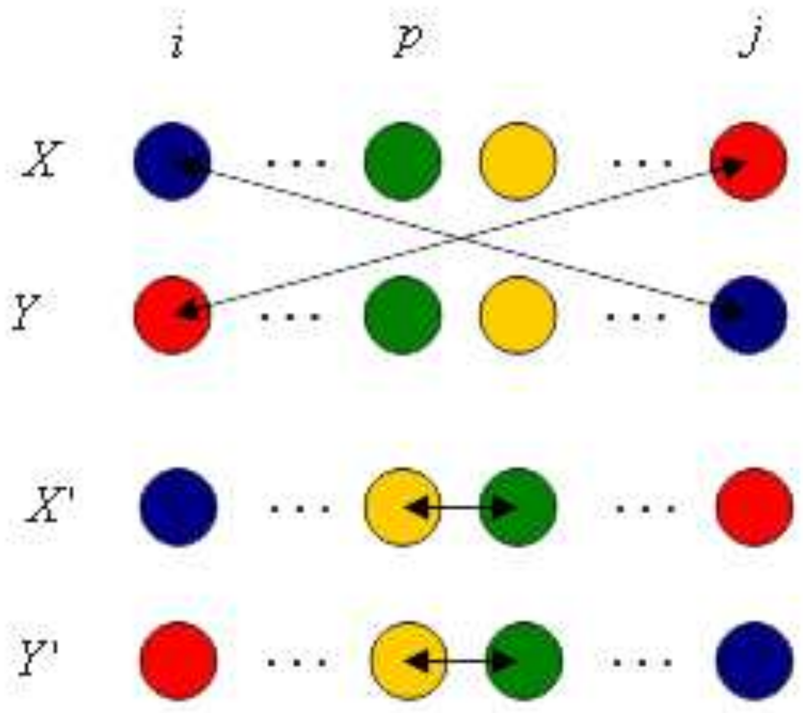
Linear Extensions - Analysis

we will show that:

- If $p = i - 1$ or $p = j$ then $\delta(X, Y)$ can increase by 1
- If $p = i$ or $p = j - 1$ then $\delta(X, Y)$ can decrease by 1
- Else $\delta(X, Y)$ will not change

We will separate the cases: $j - i = 1$ and $j - i > 1$ starting with $j - i > 1$

If $p \notin \{i-1, i, j-1, j\}$ then $\delta(X', Y') = \delta(X, Y)$

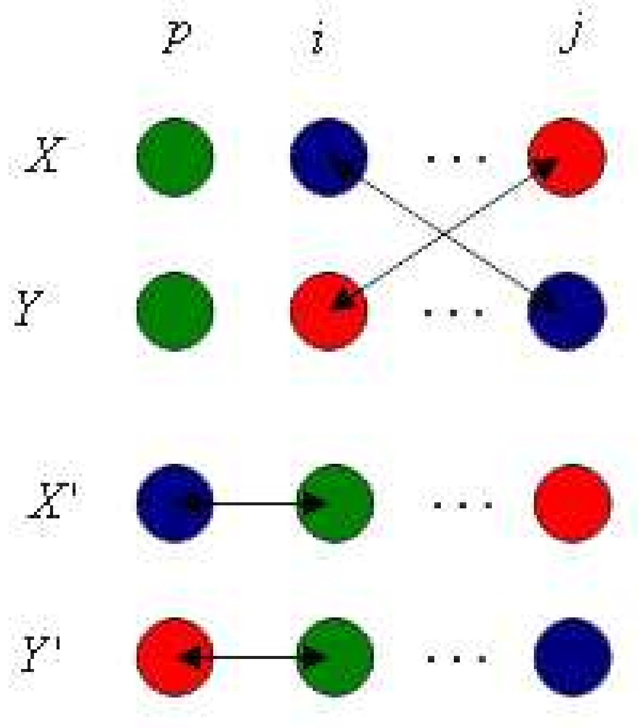


Linear Extensions - Analysis

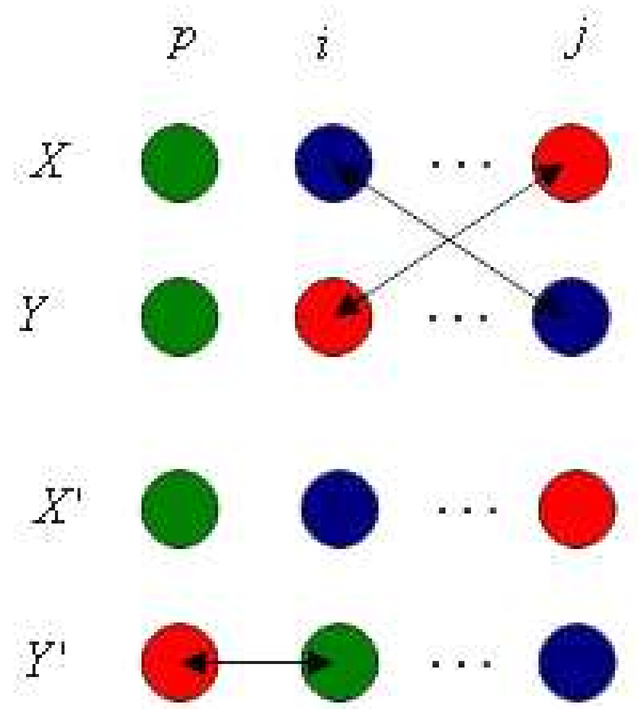
If $p = i - 1$ or $p = j$ then:

- With probability $\frac{1}{2}$ do nothing in both states
- Otherwise $\delta(X', Y') \leq \delta(X, Y) + 1$:
  - The transition is successful in both X and Y
  - The transition is successful in neither
  - The transition is successful in exactly one

If the he transition is successful in both X and Y then
$\delta(X', Y') = \delta(X, Y) + 1$

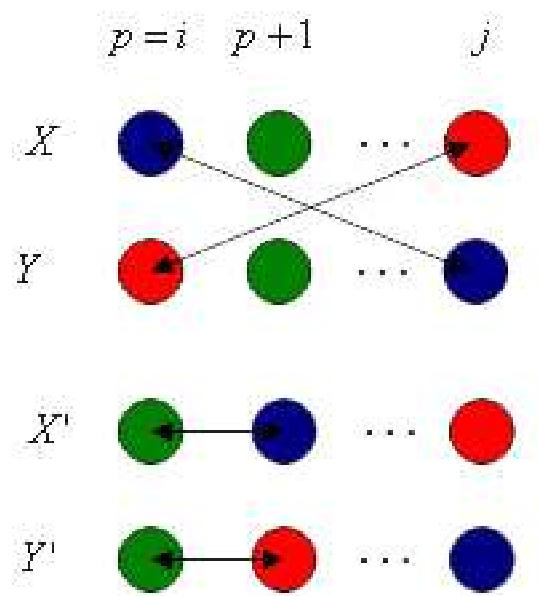


If the he transition is successful exactly one then $\delta(X', Y') = \delta(X, Y) + 1$



If $p = i$ or $p = j - 1$ then:

- With probability $\frac{1}{2}$ do nothing in both states
- Otherwise transition must be successful in both X and Y and $\delta(X', Y') = \delta(X, Y) - 1$



Linear Extensions - Analysis

For $j - i > 1$ we have:

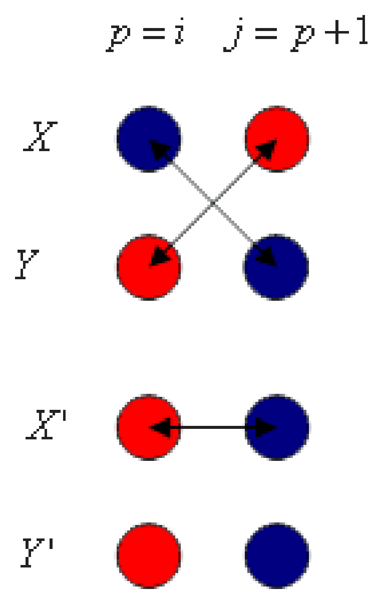
$$\mathbb{E}[\delta(X', Y')] \leq \delta(X, Y) + \frac{1}{2}\left(f(i-1) - f(i)) - f(j-1) + f(j)\right))$$

For $j - i = 1$ we have:

- If $p = i - 1$ or $p = j$, $\delta(X, Y)$ will increase by 1 with probability at most $1/2$
- If $p = i = j - 1$, $\delta(X, Y)$ will decrease by 1
- Else $\delta(X, Y)$ will not change

So we get:

$$\begin{aligned}\mathbb{E}[\delta(X',Y')] &\leq \delta(X,Y) + \frac{1}{2}\left(f(i-1) + f(j)\right) - f(i) \\ &= \delta(X,Y) + \frac{1}{2}\left(f(i-1) - f(i)) - f(j-1) + f(j)\right)\end{aligned}$$



Linear Extensions - Analysis

If we pick a concave distribution $f(i) = i(n-i)/K$ (where $K = O(n^3)$ is a normalization factor) we get:

$$\forall i < n, f(i) - f(i-i) = \frac{n+1-2i}{K}$$

and since $\delta(X,Y) = j - i$:

$$\begin{aligned}
\mathbb{E}[\delta(X',Y')] &\leq \delta(X,Y) + \frac{1}{2}\left(f(i-1) - f(i) - f(j-1) + f(j)\right) \\
&= \delta(X,Y) + \frac{1}{2K}\left((n+1-2j) - (n+1-2i)\right) \\
&= \delta(X,Y) + \frac{1}{K}(i-j) \\
&= (j-i) - \frac{1}{K}(j-i) \\
&= (1 - \frac{1}{K}) \cdot \delta(X,Y)
\end{aligned}$$

- We set β to $1 - 1/K$
- D has diameter $d = O(n^2)$
- By path coupling lemma we get $\Pr[X_t \neq Y_t] \leq \beta^t \cdot d$
- And by coupling lemma:

$$\tau(\epsilon) \leq O(K \log(d/\epsilon)) = O(n^3 \log(n \cdot \epsilon^{-1}))$$

Summary

- Intuition for path coupling
- Path coupling lemma
- A random q-coloring for $q = 2\Delta + 1$
- Path coupling for Gibbs sampler
- Linear extensions of a partial order

Questions

Reference

- Alan Freeze - Notes on counting
- Russ Bubley - Randomized Algorithms: Approximation, Generation and Counting
- Mark Jerrum - Lecture Notes. Counting, sampling and integrating: algorithms and complexity