Markov-chain Monte Carlo: A modern primer

Lecture 1: Fundamentals
Part 2/2: Single particle on the path graph

Werner Krauth
Département de physique, Ecole normale supérieure
Paris, France

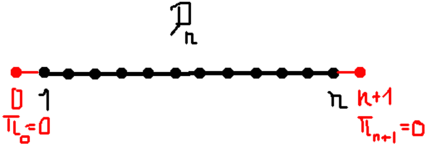
A Set of Lectures
University of Kent
Canterbury, Great Britain

14-17 November 2022

References

- F. Chen, L. Lovasz, I. Pak, **Lifting Markov Chains to Speed up Mixing.** Proceedings of the 17th Annual ACM Symposium on Theory of Computing, 275 (1999) `http://www.math.ucla.edu/~pak/papers/stoc2.pdf`
- P. Diaconis, S. Holmes, R. M. Neal, **Analysis of a nonreversible Markov chain sampler** Ann. Appl. Probab. 10, 726–752 (2000) `https://projecteuclid.org/download/pdf_1/euclid.aoap/1019487508`
- M. Hildebrand, **Rates of convergence of the Diaconis-Holmes-Neal Markov chain sampler with a V-shaped stationary probability**, Markov Proc. Rel. Fields 10, 687–704 (2004)
- W. Krauth, **Event-Chain Monte Carlo: Foundations, Applications, and Prospects**, Front. Phys. 9:663457. `https://www.frontiersin.org/article/10.3389/fphy.2021.663457`

Metropolis algorithm on path graph (1/5)

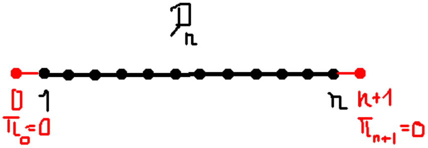


- Path graph $\mathcal{P}_n$ so that $\Omega_n = \{1, \ldots, n\}$.
- Phantom vertices and edges.

Metropolis algorithm (NB: $P_{ij} = \mathcal{A}_{ij}\mathcal{P}_{ij}$ for $i \neq j$):

1. Move set $\mathcal{L} = \{+, -\}$.
2. $\mathcal{A}$ flat $\to \sigma =$ choice($\mathcal{L}$).
3. Metropolis filter: Accept with probability $\min(1, \pi_j/\pi_i)$.

Metropolis algorithm on path graph (2/5)



- Detailed balance:

$$\underbrace{\pi_i P_{ij}}_{\mathcal{F}_{ij}} = \underbrace{\pi_j P_{ji}}_{\mathcal{F}_{ji}}$$

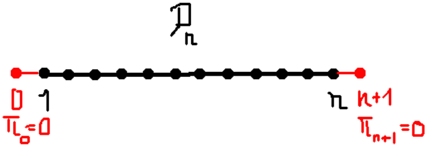
- Metropolis algorithm:

$$\mathcal{F}_{ij} = \frac{1}{2} \min\left(\pi_i, \pi_j\right) \Leftrightarrow P_{ij} = \frac{1}{2} \min\left(1, \pi_j/\pi_i\right)$$

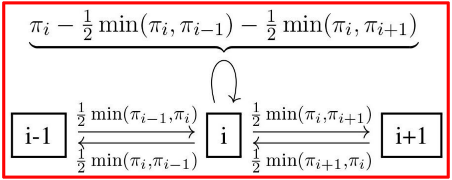
- Metropolis filter (NB: $P_{ij} = \mathcal{A}_{ij}\mathcal{P}_{ij}$):

$$\mathcal{P}_{ij} = \min\left(1, \pi_j/\pi_i\right)$$

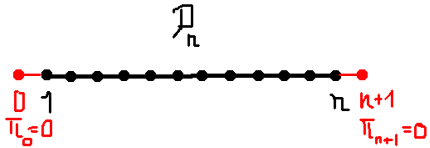
Metropolis algorithm on path graph (3/5)



- Global balance ($\pi_i = \sum_j \pi_j P_{ji} = \sum_j \mathcal{F}_{ji}$):



- Irreducibility OK if no holes in π.
- Aperiodicity OK, thanks to boundaries



procedure `metro-path`

input x

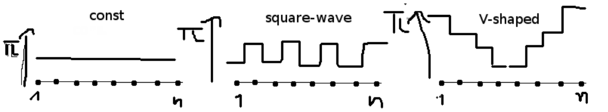
$\sigma \leftarrow$ `choice`$(\mathcal{L})$ $(\mathcal{L} = \{-1, +1\})$

if $(\text{ran}\,(0,1) < \pi_{x+\sigma}/\pi_x)$ **then**

$\quad \{\, x_i \leftarrow x_i + \sigma$

output x

———



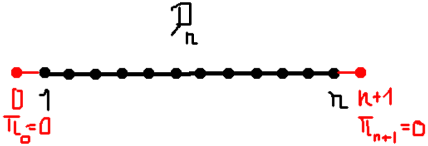
Model	Definition	Conductance	t_{mix}
Flat	$\pi_i = \frac{1}{n}$	$\mathcal{O}(1/n)$	$\mathcal{O}(n^2)$
Square	$\pi_{2k-1} = \frac{2}{3n}, \pi_{2k} = \frac{4}{3n}$	$\sim 2/(3n)$	$\mathcal{O}(n^2)$
V-shape	$\pi_i = \frac{4}{n^2}\|\frac{n+1}{2} - i\|$	$\sim 2/n^2$	$\mathcal{O}(n^2 \log n)$

NB: Graph diameter n.
NNB: π normalized.
NNNB: Bottleneck between $i = \frac{n}{2}$ and $j = \frac{n}{2} + 1$.

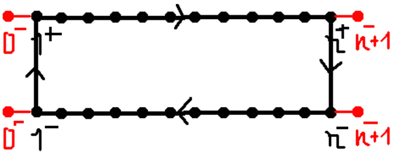
Lifting on the path graph (1/7)

Probability distribution $\pi = (1/n, \ldots, 1/n)$ (Diaconis et al. 2000)

- "Collapsed" Markov chain:



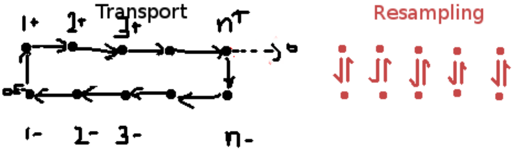
- "Lifted" Markov chain $\hat{\Omega} = \Omega \times \{-, +\}$:



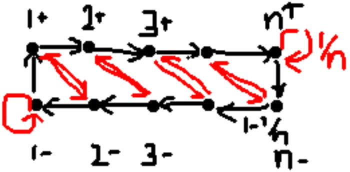
- Irreducible, but not aperiodic. $\hat{\pi}_{i,\sigma} = \frac{1}{2}\pi_i = 1/(2n)$

Probability distribution $\pi = (1/n, \ldots, 1/n)$ (Diaconis et al. 2000)

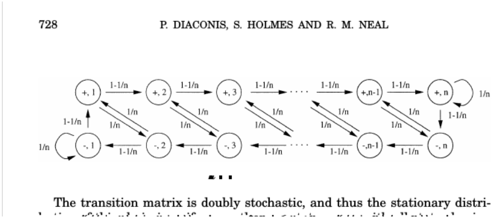
- Transport + resampling



- "Lifted" Markov chain $\hat{\Omega} = \Omega \times \{-,+\}$:



Lifting on the path graph (3/7)


728 P. DIACONIS, S. HOLMES AND R. M. NEAL

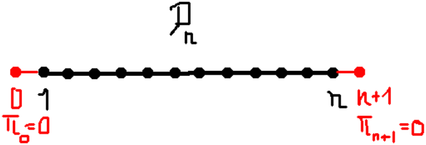
The transition matrix is doubly stochastic, and thus the stationary distri-

- Diaconis, Holmes, Neal (2000).
- Ω (samples) + $\mathcal{L}$ (moves) $\rightarrow \hat{\Omega} = \Omega \times \mathcal{L}$ (lifted samples)
- Resampling with rate $1/n$.
- Phantoms illustrate the “rejection $\rightarrow$ lifting” mystery.
- Let’s generalize.

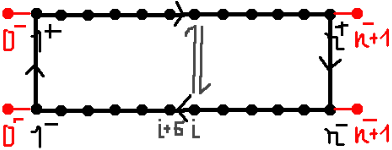
Lifting on the path graph (4/7)

General probability distribution $\pi = (\pi_1, \ldots, \pi_n)$

- "Collapsed" Markov chain:



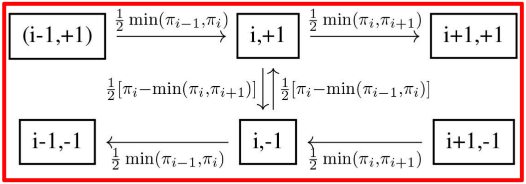
- "Lifted" Markov chain $\hat{\Omega} = \Omega \times \{-, +\}$:



- Replace all rejections by lifting moves.

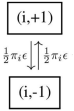
Lifting on the path graph (5/7)

- “Lifted Markov chain: Transport”



NB: The $\frac{1}{2} \Leftrightarrow \hat{\pi}_{i,\sigma} = \frac{1}{2}\pi_i$

- “Lifted Markov chain: Resampling”



- Resampling can often be dropped

Lifting on the path graph (6/7)

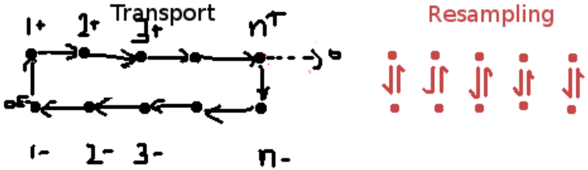
- Transport

procedure transport-path
input $\{x, \sigma\}$ (configuration $\in \hat{\Omega} = \Omega \times \{+,-\}$)
if (ran $(0,1) < \pi_{x+\sigma}/\pi_x$) **then**
  $\{ x_i \leftarrow x_i + \sigma$
else
  $\{ \sigma \leftarrow -\sigma$
output $\{x, \sigma\}$
——

- Resampling

procedure resample-path
input $\{x, \sigma\}$ (configuration $\in \hat{\Omega} = \Omega \times \{+,-\}$)
if (ran $(0,1) < p$) **then** (p: resampling rate)
  $\{ \sigma \leftarrow -\sigma$
output $\{x, \sigma\}$
——

- Mix freely!



Model	Conductance	t_{mix} (collapsed)	t_{mix} (lifted)
Flat	$\mathcal{O}(1/n)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n)$
Square	$\mathcal{O}(1/n)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$
V-shape	$\mathcal{O}(1/n^2)$	$\mathcal{O}(n^2 \log n)$	$\mathcal{O}(n^2)$