Èdouard Lucas:

The theory of recurrent sequences is an inexhaustible mine which contains all the properties of numbers; by calculating the successive terms of such sequences, decomposing them into their prime factors and seeking out by experimentation the laws of appearance and reproduction of the prime numbers, one can advance in a systematic manner the study of the properties of numbers and their application to all branches of mathematics.

A Friendly Introduction to Markov Chain Monte Carlo

Daryl DeFord

MIT/Tufts
Metric Geometry and Gerrymandering Group

Voting Rights Data Institute
June 18, 2019

Outline

Code for this talk:

Example

What is $1 + 1$?

Code for this talk:

Example

This guide follows my notes on discrete MCMC and is accompanied by a collection of 14 interactive tools that are linked here with GitHub source here. Each tool consists of an interactive terminal that allows you to vary the input values and provides many different types of plots and visualizations. This is a great opportunity to experiment with these concepts and build some extra intuition around the ideas.

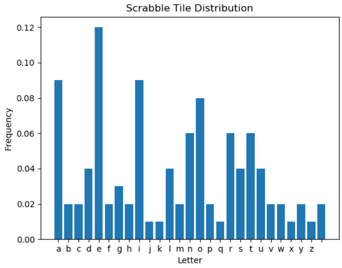
Scrabble Tiles

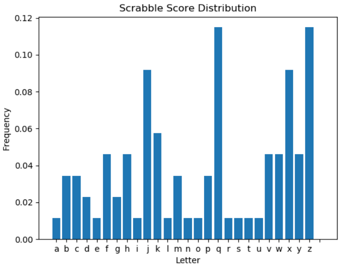
The game of Scrabble uses 100 square tiles that are drawn from a bag. Each tile is labelled with a letter (or a space) and a number, which represents the score of the tile. The number of tiles and thee score of each letter are displayed in the table below:

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Frequency	9	2	2	4	12	2	3	2	9	1	1	4	2	6
Score	1	3	3	2	1	4	2	4	1	8	5	1	3	1
Letter	O	P	Q	R	S	T	U	V	W	X	Y	Z	' '	
Frequency	8	2	1	6	4	6	4	2	2	1	2	1	2	
Score	1	3	10	1	1	1	1	4	4	8	4	10	0	

Table: Frequencies and point values of Scrabble tiles.

Discrete Distributions





Expected Values

The *expected value* of a random variable is a weighted average of the values of the state space, where the weights are given by the probabilities of the individual states.

$$\frac{9}{100}\cdot 1+\frac{2}{100}\cdot 3+\frac{2}{100}\cdot 3+\frac{4}{100}\cdot 2+\frac{12}{100}\cdot 1+\frac{2}{100}\cdot 4+\frac{3}{100}\cdot 2+\frac{2}{100}\cdot 4+\frac{9}{100}\cdot 1+\frac{1}{100}\cdot 8+\frac{1}{100}\cdot 5+\frac{4}{100}\cdot 1+\frac{2}{100}\cdot 3+\frac{6}{100}\cdot 1+$$

$$\frac{8}{100}\cdot 1+\frac{2}{100}\cdot 3+\frac{1}{100}\cdot 10+\frac{6}{100}\cdot 1+\frac{4}{100}\cdot 1+\frac{6}{100}\cdot 1+\frac{4}{100}\cdot 1+\frac{2}{100}\cdot 4+\frac{2}{100}\cdot 4+\frac{1}{100}\cdot 8+\frac{2}{100}\cdot 4+\frac{1}{100}\cdot 10+\frac{2}{100}\cdot 0=$$

$$\frac{187}{100}=1.87$$

Try it out:

- Die Rolling
- Scrabble Tiles

Try it out!

Example

Design a die with at least 8 faces that has an expected value of 6.

Example

How many rolls does it take for your expected value to get within .05.

Example

Find the smallest number of draws to get the Scrabble error under .01.

Game: Cooperative War

Rules:

1. Nominate a dealer in your group of 3 players
2. Deal 2 cards to each player
3. Each round begins with the dealer playing the highest value card from their hand.
4. The round continues counterclockwise with each player playing the highest card in their hand **only if** it is higher than the previously played card. If not, skip that player and move on to the next.
5. The round ends once each player has had an opportunity to play a card.
6. You (collectively) win if all players have played all of their cards at the end of the second round.

Cooperative War Example 1

- Hands:
  - Dealer $= \{J\spadesuit, 9\clubsuit\}$, Player 2 $=\{2\spadesuit, 3\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer:
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 1

- Hands:
  - Dealer $= \{9\clubsuit\}$, Player 2 $=\{2\spadesuit, 3\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 1

- Hands:
  - Dealer $= \{9\clubsuit\}$, Player 2 $=\{2\spadesuit, 3\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: ☹
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{J\spadesuit, 9\clubsuit\}$, Player 2 $=\{Q\spadesuit, K\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer:
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{9\clubsuit\}$, Player 2 $=\{Q\spadesuit, K\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{9\clubsuit\}$, Player 2 $=\{Q\spadesuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{9\clubsuit\}$, Player 2 $=\{Q\spadesuit\}$, Player 3 $= \{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{\}$, Player 2 $=\{Q\spadesuit\}$, Player 3 $= \{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{\}$, Player 2 $=\{\}$, Player 3 $= \{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2: $Q\spadesuit$
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{\}$, Player 2 $=\{\}$, Player 3 $= \{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2: $Q\spadesuit$
  - Player 3: ☹

Cooperative War Example 2

- Hands:
  - Dealer $= \{J\spadesuit, 9\clubsuit\}$, Player 2 $=\{Q\spadesuit, K\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2: $Q\spadesuit$
  - Player 3: ☹

Cooperative Solitaire Results?

Question

What is the probability that you win, given a randomly shuffled deck?

Cooperative Solitaire Results?

Question

What is the probability that you win, given a randomly shuffled deck?

Answer

Try it out!

Cooperative Solitaire Results?

Question

What is the probability that you win, given a randomly shuffled deck?

Answer

Try it out!

Answer

Simulate! https://people.csail.mit.edu/ddeford/solitaire

War Solitaire

Rules:

1. Fix n and k
2. Lay the top n cards of the deck on the table in order $c_1, c_2, \ldots, c_n$
3. You lose unless $c_1 < c_2 < \cdots < c_n$
4. Repeat k times
5. You win if you haven't lost.

War Solitaire Results?

Question

What is the probability that you win, given a randomly shuffled deck?

War Solitaire Results?

Question

What is the probability that you win, given a randomly shuffled deck?

Answer

Try it out!

War Solitaire Results?

Question

What is the probability that you win, given a randomly shuffled deck?

Answer

Try it out!

Answer

Simulate! https://people.csail.mit.edu/ddeford/war

Geometric Probability

Question

What is the expected distance between two random points on $[0, 1]$*?*

Geometric Probability

Question

What is the expected distance between two random points on $[0,1]$*?*

Answer

$$\int_0^1 \int_0^1 |x-y| dx dy = \frac{1}{3}$$

Geometric Probability

Question

What is the expected distance between two random points on $[0,1]$*?*

Answer

$$\int_0^1 \int_0^1 |x-y| dx dy = \frac{1}{3}$$

Question

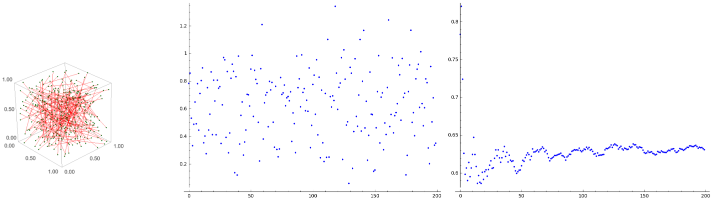
What is the expected distance between two random points on $[0,1]^n$*?*

Answer

$$\int_0^1 \cdots \int_0^1 \sqrt{\sum_{j=1}^n (x_j - y_j)^2} dx_1 \cdots dx_n dy_1 \cdots dy_n = \quad ☹$$

Random Points

Try it out: Cube Distances

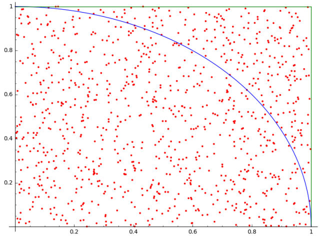


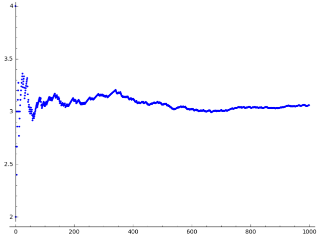
Numerical Integration

Question

What is the area "under" the curve?

$$\int_0^1 \int_0^{\sqrt{1-x^2}} 1 dx dy$$





Properties of Monte Carlo Methods

- Draw (independent) samples from a random distribution
- Compute some measure for each draw
- Repeat lots and lots of times
- Average/aggregate the derived data

Properties of Monte Carlo Methods

- Draw (independent) samples from a random distribution
- Compute some measure for each draw
- Repeat lots and lots of times
- Average/aggregate the derived data

Today, we tend to take access to random numbers for granted but some of the first applications of Monte Carlo were physical systems for generating random numbers.

What is a Markov chain?

Definition (Markov Chain)

A sequence of random variables $X_1, X_2, \ldots,$ is called a Markov Chain if

$$\mathbb{P}(X_n = x_n : X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = x_n : X_{n-1} = x_{n-1}).$$

What is a Markov chain?

Definition (Markov Chain)

A sequence of random variables $X_1, X_2, \ldots,$ is called a Markov Chain if

$$\mathbb{P}(X_n = x_n : X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = x_n : X_{n-1} = x_{n-1}).$$

Definition (Transition Probability)

Given a finite state space $X = x_1, x_2, \ldots, x_n$ we can specify a Markov chain over X with transition probabilities $p_{i,j} = \mathbb{P}(X_m = i : X_{m-1} = j)$ and associated transition matrix $P = [p_{i,j}]$.

Random Walks

Definition (Random Walk)

We can also view a Markov chain as random walk on a directed, weighted graph, with weights given by the $p_{i,j}$.

For any pair of letters α and β, we can compute the probability that the ant steps from α to β by:

$$\mathbb{P}(\alpha \mapsto \beta) = \begin{cases} \dfrac{1}{\#\text{neighbors of } \alpha} & \text{if } \alpha \text{ is next to } \beta \\ 0 & \text{otherwise.} \end{cases}$$

Letter Examples

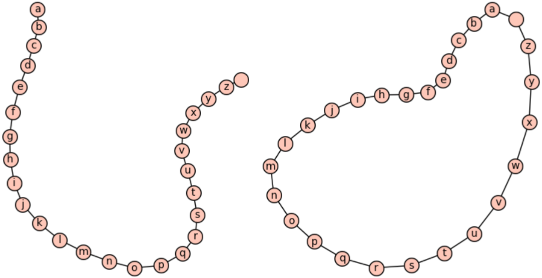
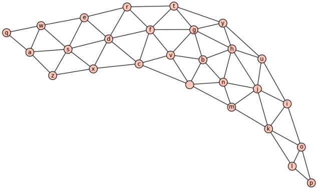
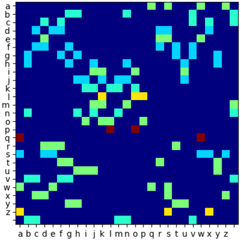


Figure: The alphabet path and cycle graphs.

Keyboard





Animation link

Letter Walk Examples

- Path
- Cycle
- Keyboard
- i.i.d. Scrabble
- i.i.d. Uniform

Try it out: Markov Chain Samples

Try it Out!

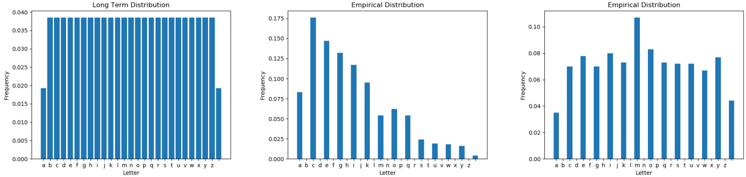
Example

Which of the 5 Markov chains are closest to their stationary distribution after 50 steps? 1000 steps?

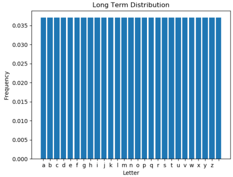
Example

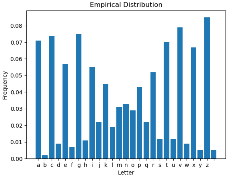
Which combination of walk and starting letter leads to the most attractive distribution plot?

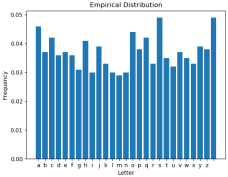
Path Walk


Long Term Distribution
Frequency
Letter
Empirical Distribution
Frequency
Letter
Empirical Distribution
Frequency
Letter

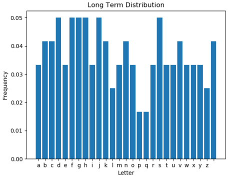
Cycle Walk

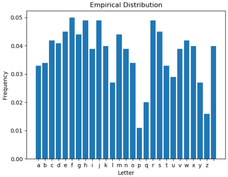

Long Term Distribution
Frequency
Letter
a b c d e f g h i j k l m n o p q r s t u v w x y z
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035

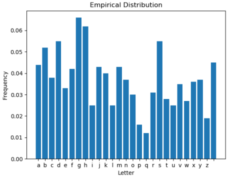

Empirical Distribution
Frequency
Letter
a b c d e f g h i j k l m n o p q r s t u v w x y z
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08


Empirical Distribution
Frequency
Letter
a b c d e f g h i j k l m n o p q r s t u v w x y z
0.00
0.01
0.02
0.03
0.04
0.05

Keyboard Walk







Desirable Adjectives

- Irreducible: A chain is irreducible if each state is (eventually) reachable from every other state.
- Aperiodic: A chain is aperiodic if for each state, the GCD of the lengths of the loops, starting and ending at that state is equal to 1.
- Steady State Distribution: A distribution π is said to be a steady state of the chain if $\pi = \pi P$. For simple random walks on graphs this is proportional to the degree of each node.
- Reversible: A chain with steady state π is called reversible if it satisfies the detailed balance condition:

$$\pi_i P_{i,j} = \pi_j P_{j,i}$$

Key Theorem

If the chain is irreducible and aperiodic then $\lim_{m \to \infty} P^m = 1\pi$ for a unique π. Even better, if $y_1, y_2, \ldots, y_m$ are samples from π then,

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} f(y_i) = \mathbb{E}[f]$$

The key idea of MCMC is to create an irreducible, aperiodic Markov chain whose steady state distribution π is the distribution we are trying to sample from.

Score Function

- Uniform
- Scrabble Points
- Scrabble Counts
- Alphabetical (a=1, b=2, etc.)
- Vowels (consonants = 1, vowels = 100, y = 50)

Try it out: Compute Expected Values

Try it out!

Example

Using the Keyboard walk, which score function converges to the right expected value fastest.

Example

For the path walk and alphabetical score, which starting point converges most slowly?

Example

For the keyboard walk and vowel score, which starting point converges fastest?

Estimating Expected Values

Walk	Score	Actual Value	2k steps	10k steps	50k steps	100k steps
Keyboard	Points	3.075	3.129 (1.7%)	3.067 (0.3%)	3.089 (0.4%)	3.072 (0.1%)
Keyboard	Count	3.633	3.733 (2.7%)	3.587 (1.3%)	3.644 (0.3%)	3.638 (0.1%)
Keyboard	Alphabet	13.292	13.12 (1.2%)	13.34 (0.4%)	13.32 (0.2%)	13.30 (0.05%)
Keyboard	Vowels	19.13	21.02 (9.9%)	19.36 (1.2%)	19.53 (2.1%)	18.91 (1.2%)
Cycle	Points	3.22	3.17 (1.7%)	3.18 (1.2%)	3.19 (0.9%)	3.20 (0.7%)
Cycle	Count	3.70	4.00 (8%)	3.88 (4.9%)	3.71 (0.3%)	3.71 (0.2%)
Cycle	Alphabet	14	14.5 (3.5%)	14.32 (2.3%)	14.1 (0.7%)	13.88 (0.8%)
Cycle	Vowels	21.15	21.36 (1.0%)	21.76 (2.9%)	20.97 (0.8%)	21.22 (0.4%)
Path	Points	3.327	3.67 (10.3%)	3.18 (4.5%)	3.33 (0.02%)	3.31 (0.4%)
Path	Count	3.635	3.85 (5.9%)	3.68 (1.2%)	3.69 (1.6%)	3.59 (1.0%)
Path	Alphabet	14	15.75 (12.5%)	13.99 (0.07%)	14.04 (0.3%)	14.07 (0.5%)
Path	Vowels	20.02	18.69 (6.6%)	19.70 (1.6%)	19.29 (3.6%)	20.03 (0.07%)

Mixing Times

Definition (Mixing Time)

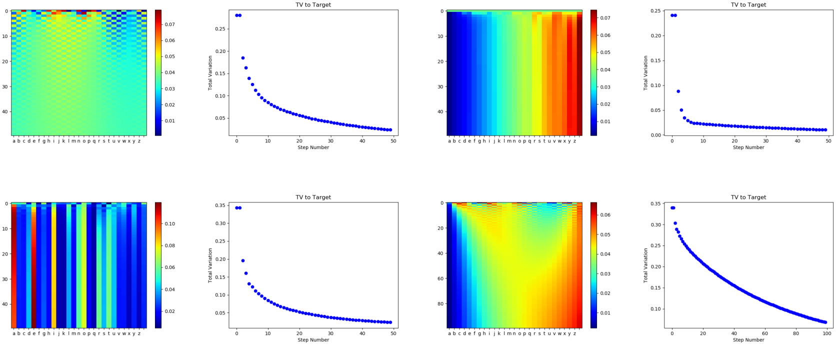
Given a Markov chain with transition matrix P, the mixing time is the smallest integer k such that

$$||\pi - QP^k||_{TV} < \frac{1}{4}$$

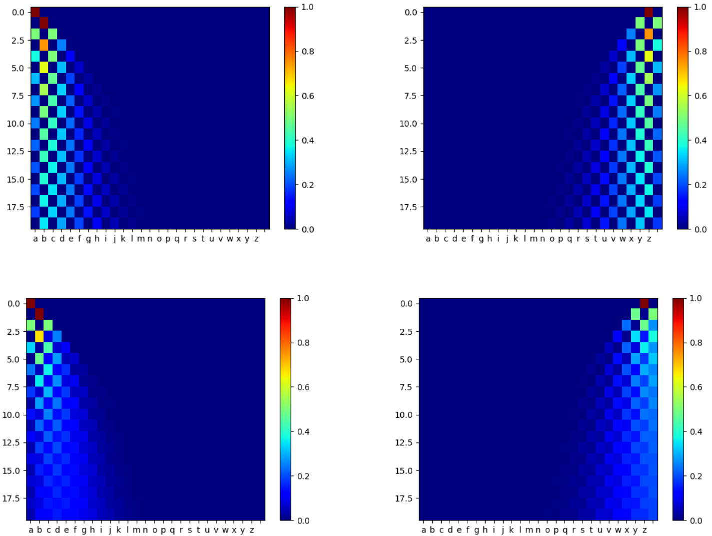
for any initial distribution Q.

Try it out: Mixing Comparison

Mixing Examples



Slow Examples



Example

Are there any pure states that converge faster than uniform states?

Example

Can you find a random state that requires more than 5 steps to get below $\frac{1}{4}$?

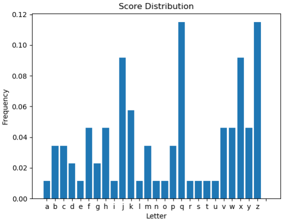
Example

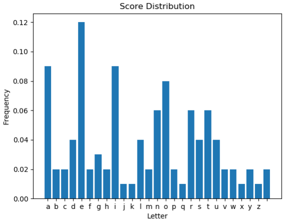
Using the path walk, which starting letter mixes fastest to the vowel distribution?

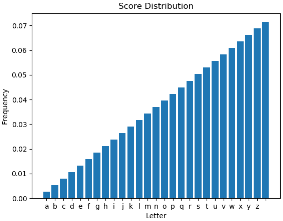
What is MCMC?

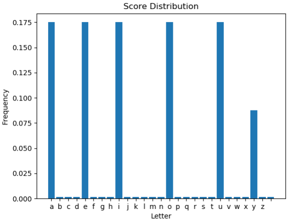
In our Monte Carlo methods we just required that we sample from our space uniformly but this isn't always easy to do. MCMC gives us a way to sample from a desired pre–defined distribution by forming a related Markov chain (or walk) over our state space, with transition probabilities determined by a multiple of the distribution that we are trying to sample from.

Score Distributions









Proportional to a distribution!?!

A common way this arises is when we have a score function or a ranking on our state space and want to draw proportionally to these scores. Given a score $s : X \to \mathbb{R}$ we want to sample from the distribution where the states appear proportional to s. That is, element $y \in X$ should appear with probability

$$\mathbb{P}(y) = \frac{s(y)}{\sum_{x \in X} s(x)}.$$

When $|X|$ is enormous, we don't want to/can't compute the denominator directly. Also, uniform sampling over–prioritizes low score subspaces.

How does it work?

Notice that we can compute ratios of probabilities, since the denominators cancel:

$$\frac{\mathbb{P}(z)}{\mathbb{P}(y)} = \frac{\frac{s(z)}{\sum_{x \in X} s(x)}}{\frac{s(y)}{\sum_{x \in X} s(x)}} = \frac{s(z)}{s(y)}.$$

This is the trick that turns out to allow us to draw samples according to s without having to compute the denominator directly.

Terminology

Score: A function $s : X \to \mathbb{R}_{\geq 0}$ that determines our target distribution.

Terminology

Score: A function $s : X \to \mathbb{R}_{\geq 0}$ that determines our target distribution.

Proposal Distribution: A Markov chain Q over X with the property that $Q(x_j : x_i)$.

Terminology

Score: A function $s : X \to \mathbb{R}_{\geq 0}$ that determines our target distribution.

Proposal Distribution: A Markov chain Q over X with the property that $Q(x_j : x_i)$.

Metric: Another function $f : X \to \mathbb{R}$ that is our quantity of interest for the distribution.

Metropolis Procedure

Given that we have a given score function, proposal distribution, metric, and initial state X_0, at each step of the Metropolis–Hastings chain $X_1, X_2, \ldots$ we follow this sequence of steps, assuming that we are currently at state $X_k = y$:

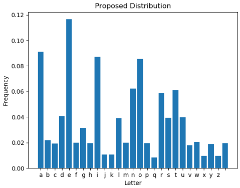
1. Generating a proposed state $\hat{y}$ according to $Q_{y,\hat{y}}$.
2. Compute the acceptance probability:

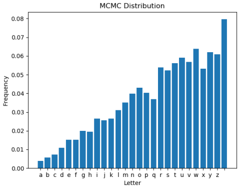
$$\alpha = \min\left(1, \frac{s(\hat{y})}{s(y)} \frac{Q_{\hat{y},y}}{Q_{y,\hat{y}}}\right)$$

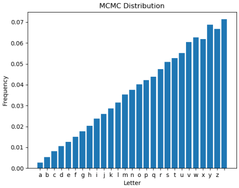
3. Pick a number β uniformly on $[0, 1]$
4. Set

$$X_{k+1} = \begin{cases} \hat{y} & \text{if } \beta < \alpha \\ y & \text{otherwise.} \end{cases}$$

Letter Example







Try it out: here

Example: Step 1

1. We uniformly pick a key on the keyboard next to 'a': 'q'.
2. We next need to compute some numbers in order to compute the acceptance probability:
   - $s(a) = 1$
   - $s(q) = 10$
   - $Q_{q,a} = \frac{1}{2}$ since 'q' has two neighbors
   - $Q_{a,q} = \frac{1}{3}$ since 'a' has three neighbors

   These let us compute:

$$\alpha = \min(1, \frac{10}{1}\frac{\frac{1}{2}}{\frac{1}{3}}) = \min(1, 15) = 1$$

3. Uniformly pick $\beta = .188256$
4. Set the next state to be 'q' since $\beta < \alpha$.

Example: Step 2

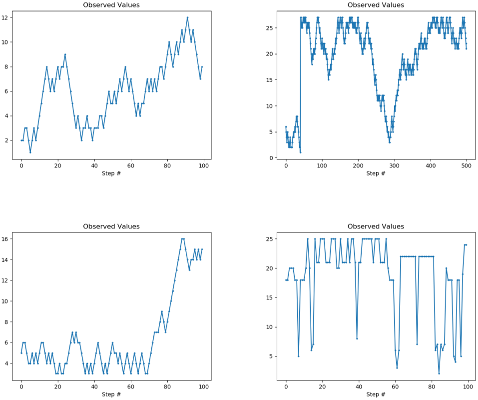
1. We uniformly pick a neighbor of ‘q’ and get ‘w’.
2. We again need to compute some numbers:
   - $s(q) = 10$
   - $s(w) = 4$
   - $Q_{q,w} = \frac{1}{2}$ since ‘q’ has two neighbors
   - $Q_{w,q} = \frac{1}{3}$ since ‘w’ has four neighbors

   These let us compute:

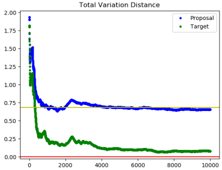
$$\alpha = \min(1, \frac{4}{10}\frac{\frac{1}{4}}{\frac{1}{2}}) = \min(1, .2) = .2$$

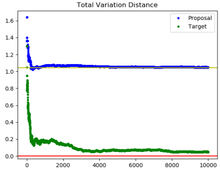
3. Uniformly pick $\beta = .7593544$
4. Set the next state to be ‘q’ since $\beta > \alpha$.

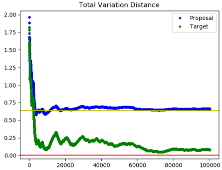
Traces

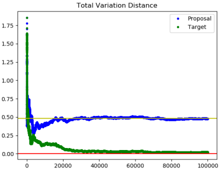

Observed Values
Step #
Observed Values
Step #
Observed Values
Step #
Observed Values
Step #

Total Variation



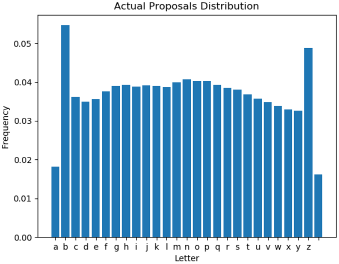


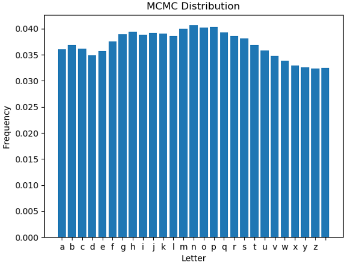




Simplifications

- Uniform score
- Equal transitions
- How to transform the cycle to the path





Try it out!

Example

Try a chain with each combination of proposal and score function. Which ones work best?

MCMC on Partitions

MCMC on Districts

- State space: Permissible Partitions
  - Contiguity
  - Population Imbalance
  - Splitting Rules
- Proposal
  - Change Assignments
  - Rejection Sampling
- Score Function
  - Prioritize rule compliance
  - Uniform

MCMC on Districts

- **State space: Permissible Partitions**
  - **Contiguity**
  - **Population Imbalance**
  - **Splitting Rules**
- Proposal
  - Change Assignments
  - Rejection Sampling
- Score Function
  - Prioritize rule compliance
  - Uniform

Permissible Partitions

Problems:

- Metagraph
  - There are too many plans
  - Many of the plans are undesirable
  - There are too many adjacent states
  - We aren't going to visit most neighbors
  - Determining permissibility is expensive
- How to operationalize the state rules

Solutions:

- Rejection Sampling
  - At each step we generate proposals until we find one that is in the state space
- Try lots of options

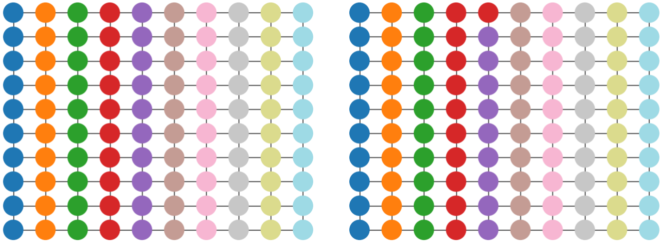
MCMC on Districts

- State space: Permissible Partitions
  - Contiguity
  - Population Imbalance
  - Splitting Rules
- **Proposal**
  - **Change Assignments**
  - **Rejection Sampling**
- Score Function
  - Prioritize rule compliance
  - Uniform

Proposal

Instead of trying to define our transitions directly as related to some distribution, we instead define stochastic rules for changing the assignments of some number of nodes in the current state.

Flip Proposal



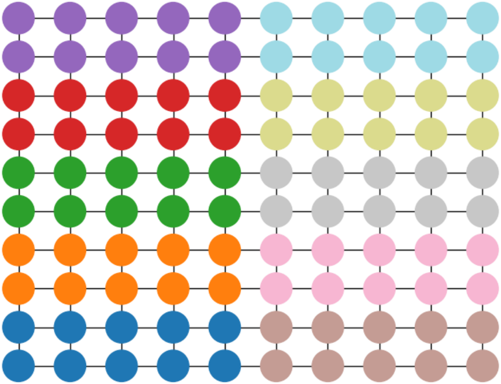
Recombination Steps

1. At each step, select two adjacent **districts**
2. Merge the subunits of those two districts
3. Draw a new district using **your favorite** bipartitioning method.
4. Repeat

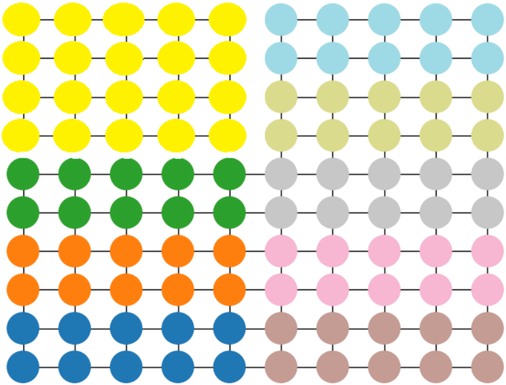
Recombination Steps

1. At each step, select two adjacent **districts**
2. Merge the subunits of those two districts
3. Draw a spanning tree for the new super–district
4. Delete an edge leaving two population balanced districts
5. Repeat
6. (Optional) Mix with single edge flips

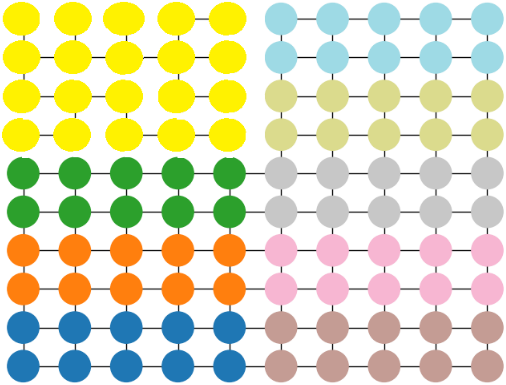
Recombination Step Example



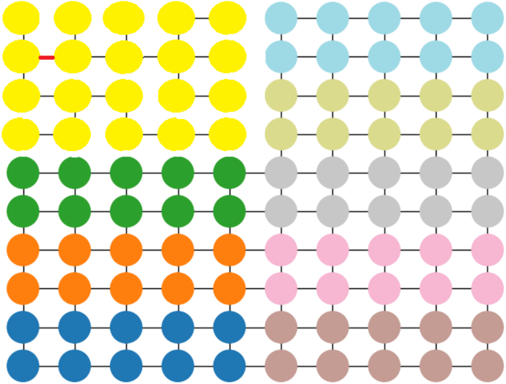
Recombination Step Example



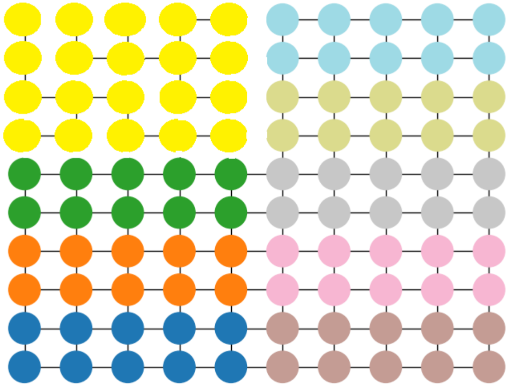
Recombination Step Example



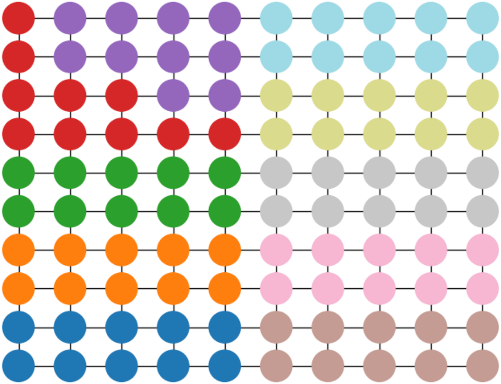
Recombination Step Example



Recombination Step Example



Recombination Step Example



MCMC on Districts

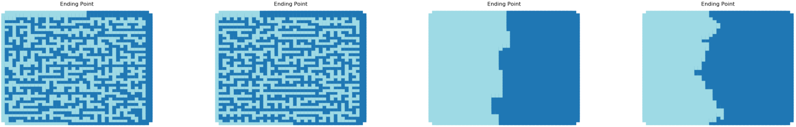
- State space: Permissible Partitions
  - Contiguity
  - Population Imbalance
  - Splitting Rules
- Proposal
  - Change Assignments
  - Rejection Sampling
- **Score Function**
  - **Prioritize rule compliance**
  - **Uniform**

Score Functions

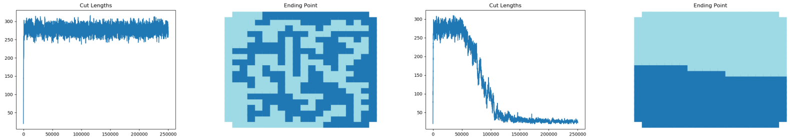
Motivated by approaches used in statistical physics we can construct an "energy" function $s(P) = e^{-\beta \sum \alpha_i \sigma_i(P)}$ where P is a partition of our graph, β is a constant called the inverse temperature and the α_i are linear coefficients on the σ_i which are the individual score functions representing our metrics of interest. We then want to sample partitions proportional to their score $\mathbb{P}(P) = \dfrac{f(P)}{\sum f(Q)}$. The denominator here is a sum over all partitions of the graph that satisfy our binary constraints. Since we are unable to compute this set on real world examples this is a perfect setting for applying the MCMC method.

Cut Edges

Given a partition $P = (A, B)$ we set $\sigma_{pop}((A, B)) = ||A| - |B||$ to be the difference in the sizes of the partitions and our score function is $e^{-\sigma_{pop}(A,B)}$. Sampling proportional to this distribution means that a plan with exactly balanced populations should appear approximately $e^{10} \sim 22,026$ as frequently as a plan with 10 more nodes in A than B.

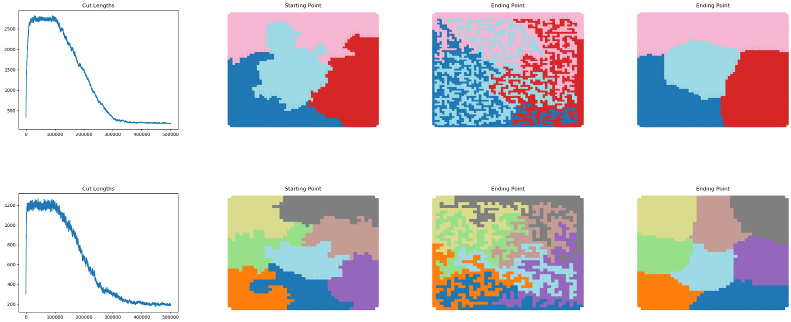


Annealing


Cut Lengths
300
250
200
150
100
50
0
50000
100000
150000
200000
250000
Ending Point
Cut Lengths
300
250
200
150
100
50
0
50000
100000
150000
200000
250000
Ending Point

(d)

More Annealing



Creativity!

- Scores
  - Compactness
  - Aggregating measures
  - Hard vs. Soft constraints
  - ...
- Proposals
  - Boundary Flip
  - Tree Methods
  - Your favorite graph method here
  - ...
- Metrics
  - Sorted vote percentage vector
  - Partisan Metrics
  - Competitiveness?
  - ...

The End

Thanks!