Reversible-jump Markov chain Monte Carlo

Andrea Gesmundo @ University of Geneva
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo

Andrey Markov
Reversible-jump Markov chain Monte Carlo

Andrey Markov
Reversible-jump Markov chain Monte Carlo

DIRECTED GRAPH
Reversible-jump Markov chain Monte Carlo

STATES

- E
- A
Reversible-jump Markov chain Monte Carlo

STATES

TRANSITIONS

E

A

0.3

0.7

0.4

0.6
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo

STATES

TRANSITIONS

PROBABILITIES

Markov chain IS PATH IN THE GRAPH:
Reversible-jump Markov chain Monte Carlo

Markov chain IS PATH IN THE GRAPH: E
Reversible-jump Markov chain Monte Carlo

Markov chain IS PATH IN THE GRAPH: E → A → E
Reversible-jump Markov chain Monte Carlo

STATES

TRANSITIONS

PROBABILITIES

Markov chain IS PATH IN THE GRAPH: 

E E
Reversible-jump Markov chain Monte Carlo

Markov chain IS PATH IN THE GRAPH: E E E A
Reversible-jump Markov chain Monte Carlo

Markov chain IS PATH IN THE GRAPH:
Reversible-jump Markov chain Monte Carlo

STATES

TRANSITIONS

PROBABILITIES

PROBABILITY OF PATH = \prod PROBABILITIES OF TRANSITIONS

Markov chain IS PATH IN THE GRAPH: E E A E
Reversible-jump Markov chain Monte Carlo

- **STATES**
- **TRANSITIONS**
- **PROBABILITIES**

MEMORYLESS!
Reversible-jump Markov chain Monte Carlo

MEMORYLESS!

Transition probabilities depend only on source and target states
Reversible-jump Markov chain Monte Carlo

Manhattan example: The distance to where you are going depends only on where you are. Does not depend on where you came from.
Reversible-jump **Markov chain** Monte Carlo

**Formal definition:**

**Directed Graph:**
States & Edges: $S$, $V$
Transition probabilities: $P(x \mid x')$, $x, x' \in S$
Reversible-jump Markov chain Monte Carlo

Formal definition:

Directed Graph:
- States & Edges: $S$, $V$
- Transition probabilities: $P(x \mid x')$, $x, x' \in S$

Markov chain:
- Is a sequence of random variables: $X_1, X_2, X_3,$
- With independence Markov property:

$$\Pr(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \Pr(X_{n+1} = x \mid X_n = x_n).$$
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo

Monte Carlo: is a city-state between France and Italy with a famous Casino
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!

Here we refers to a Sampling technique: where is the connection?
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!

Stanislaw Ulam coined this name while working on the Manhattan project
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!

MEANWHILE in MONTE CARLO

Stanislaw Ulam's uncle gambled away his money @casino
Reversible-jump Markov chain Monte Carlo

**SAMPLING TECHNIQUE?**

MEANWHILE in MONTE CARLO

*Stanislaw Ulam’s uncle* gambled away his money @casino
Reversible-jump Markov chain Monte Carlo SAMPLING TECHNIQUE?
Problem:
You want to find the area under the red line
Quarter of circle of ray =1
Problem:
You want to find the area under the red line
Quarter of circle of ray = 1

Option 1, Analytical solution:
\[
\text{quarter\_circle\_area} = \frac{\pi r^2}{4}
\]
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE?

Problem:
You want to find the area under the red line
Quarter of circle of ray = 1

Option 1, Analytical solution:
quarter_circle_area = \frac{\pi r^2}{4}

Option 2, SAMPLING!
• Draw N random points with uniform distribution
• Count the points that fall under the red line = C
• While N grows, C/N converges to quarter_circle_area
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!

Sampling Technique: Takes samples
Reversible-jump Markov chain Monte Carlo SAMPLING TECHNIQUE!

Sampling Technique: Takes samples
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!

Sampling Technique: Takes samples
Reversible-jump Markov chain Monte Carlo SAMPLING TECHNIQUE!

Sampling Technique: Takes samples
Reversible-jump Markov chain Monte Carlo

WHY SAMPLING TECHNIQUE?
Reversible-jump Markov chain Monte Carlo

WHY SAMPLING TECHNIQUE?

To find the area under the line

To find the integral of the function that describes the line
Reversible-jump Markov chain Monte Carlo

WHY SAMPLING TECHNIQUE?
To find the area under the line

\[ \int_{a}^{b} f(x) \, dx \]

To find the integral of the function that describes the line
Reversible-jump Markov chain Monte Carlo

WHY SAMPLING TECHNIQUE?

To find the area under the line

\[ \int_{a}^{b} f(x) \, dx \]

To find the integral of the function that describes the line

Draw a box!

Area of the box is easy to find
Reversible-jump Markov chain Monte Carlo

WHY SAMPLING TECHNIQUE?
To find the area under the line

To find the integral of the function that describes the line

Take samples!
To find proportion of 2 areas
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!
Reversible-jump Markov chain Monte Carlo SAMPLING TECHNIQUE!
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!
Reversible-jump Markov chain Monte Carlo

SAMPLING TECHNIQUE!

Can approximate area/integral

Can also approximate shape of the curve to find peaks
WHY SAMPLING TECHNIQUE?

Need to solve an integral?

Option 1, Analytical solution:
- Closed form solutions are not always available
+ Exact result
WHY SAMPLING TECHNIQUE?

Need to solve an integral?

Option 1, Analytical solution:
- Closed form solutions are not always available
+ Exact result

Option 2, SAMPLING!
- Approximated result
+ Always possible
+ Can balance accuracy with computing time
Sampling techniques are applied for problems that require solving complex argmax or integrals with no closed form solution.
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_{a}^{b} g(x) f(x) \, dx \]
Reversible-jump Markov chain Monte Carlo

Formal definition: Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_{a}^{b} g(x) f(x) \, dx \]

Bayesian models often require solving this kind of problems:

Predictions involve marginalization e.g.

\[ p(y|X) = \int p(y|\theta)p(\theta|X) \, d\theta \]
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_{a}^{b} g(x) f(x) \, dx \]

Conditions to apply MC:

\[ f(x) \geq 0, \quad x \in (a, b) \quad \text{Function } \geq 0 \text{ in interval} \]

\[ \int_{a}^{b} f(x) \, dx = M < \infty \quad \text{Area under Function is finite} \]
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_{a}^{b} g(x) f(x) \, dx \]

Let:
\[ f^*(x) = \frac{f(x)}{M} \]

\( f^* \) is Probability Density Function over \((a,b)\)
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_a^b g(x) f(x) \, dx \]

Considering \( f^* \) as a PDF allows us to rewrite the integral as the expected value of \( g(x) \) over \( f^* \)

\[ y = \int_a^b M g(x) f^*(x) \, dx = M \mathbb{E}_{f^*}(g(x)) \]
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_a^b g(x) f(x) \, dx \]

We can draw samples \((x_1, x_2, \ldots, x_n)\) from \(f^*\) and approximate the integral with:

\[ y \approx M \frac{1}{n} \sum_{i=1}^n g(x_i) \]
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_a^b g(x) f(x) \, dx \]

We can draw samples \((x_1, x_2, \ldots, x_n)\) from \(f^*\)
And approximate the integral with:

\[ y \approx M \frac{1}{n} \sum_{i=1}^{n} g(x_i) \]

Can you find 2 ways to find area using MC formalism?
Reversible-jump Markov chain Monte Carlo

Formal definition:
Monte Carlo is a generic framework for sampling

Problem:
\[ y = \int_{a}^{b} g(x) f(x) \, dx \]

We can draw samples \((x_1, x_2, \ldots, x_n)\) from \(f^*\) and approximate the integral with:

\[ y \approx M \frac{1}{n} \sum_{i=1}^{n} g(x_i) \]

Can you find 2 ways to find area using MC formalism?

Note that still need to solve one integral to compute M
Reversible-jump Markov chain Monte Carlo

Monte Carlo is a generic framework for sampling techniques
Reversible-jump Markov chain Monte Carlo

Monte Carlo is a generic framework for sampling techniques

Acceptance Rejection Sampling is a MC specific instance
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

Used when is not easy to draw sample from $f(x)$
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

Used when is not easy to draw sample from $f(x)$

Sample from an easy distribution $h(x)$

Called proposal distribution
Reversible-jump Markov chain Monte Carlo

**Acceptance Rejection Sampling**

Used when it is not easy to draw sample from $f(x)$

Sample from an easy distribution $h(x)$

Called proposal distribution

$h(x)$ must satisfy the condition:

There exists a constant $c > 0$ such that $f(x) \leq ch(x)$

For any $x$ in $f(x)$ domain
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

Used when is not easy to draw sample from $f(x)$

Sample from an easy distribution $h(x)$
Called proposal distribution

$h(x)$ must satisfy the condition:

There exists a constant $c > 0$ such that $f(x) \leq ch(x)$
For any $x$ in $f(x)$ domain

Draw a box!

Recurring idea of creating an envelope around function to integrate
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

During sampling we draw a sample $x^*$ from $h(x)$
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

During sampling we draw a sample \( x^\ast \) from \( h(x) \)

\( x^\ast \) is accept with probability:

\[
\alpha(x^\ast) = \frac{f(x^\ast)}{ch(x^\ast)}
\]
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

During sampling we draw a sample $x^*$ from $h(x)$

$x^*$ is accept with probability:

$$\alpha(x^*) = \frac{f(x^*)}{ch(x^*)}$$

$\alpha(x^*)$ is always <1 because of the condition on $h(x)$
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

Sampling loop:

Draw a sample $x^*$ from $h(x)$

Compute acceptance probability $\alpha(x^*)$

Draw a random number $u$ from $\text{uniform}(0, 1)$

Accept $x^*$ if $u < \alpha(x^*)$
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

\[ \alpha_1 \]

\[ \alpha_2 \]
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

A closer $h(x)$ to $f(x)$
leads to a higher acceptance rate
and more efficient sampling
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

A closer $h(x)$ to $f(x)$ leads to a higher acceptance rate and more efficient sampling.

No need to solve the integral to compute the normalizing factor $M$. 
Reversible-jump Markov chain Monte Carlo

Acceptance Rejection Sampling

A closer $h(x)$ to $f(x)$
leads to a higher acceptance rate
and more efficient sampling

No need to solve the integral to compute the
normalizing factor $M$

Can you solve the quarter_circle_area problem using AR sampling?
Reversible-jump Markov chain Monte Carlo

\[
\int \frac{4x^2 + 2x + 12}{(x^2 + 2)(x - 3)} \, dx = \int \frac{-10}{11} \frac{x + 8}{x^2 + 2} \, dx + \frac{54}{11} \int \frac{1}{x - 3} \, dx = \\
= -\frac{2}{11} \int \frac{5x + 4}{x^2 + 2} \, dx + \frac{54}{11} \int \frac{1}{x - 3} \, dx
\]

\[
I(x) = I_0 \left( \left| \int_{-\infty}^{x} \cos \left( \frac{\pi u^2}{L \lambda} \right) \, du \right|^2 + \left| \int_{-\infty}^{x} \sin \left( \frac{\pi u^2}{L \lambda} \right) \, du \right|^2 \right)
\]

SOLVE COMPLICATED INTEGRALS!!!
Reversible-jump Markov chain Monte Carlo

\[
\int \frac{4x^2 + 2x + 12}{(x^2 + 2)(x - 3)} \, dx = \int \frac{-10}{11} \frac{x + \frac{8}{11}}{x^2 + 2} \, dx + \frac{54}{11} \int \frac{1}{x - 3} \, dx = -\frac{2}{11} \int \frac{5x + 4}{x^2 + 2} \, dx + \frac{54}{11} \int \frac{1}{x - 3} \, dx
\]

\[
l(x) = I_0 \left( \left| \int_{-\infty}^{x} \cos \left( \frac{\pi u^2}{L \lambda} \right) \, du \right|^2 + \left| \int_{-\infty}^{x} \sin \left( \frac{\pi u^2}{L \lambda} \right) \, du \right|^2 \right)
\]

SOLVE COMPLICATED INTEGRALS!!!
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo

Where is the connection?
Reversible-jump Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC)

Is a sub class of MC sampling techniques

Which draws samples by creating a Markov chain
Reversible-jump Markov chain Monte Carlo

In synthesis:

There is a different $h(x)$ associated to each possible sample
Reversible-jump Markov chain Monte Carlo

In synthesis:

There is a different $h(x)$ associated to each possible sample.

After a new sample is drawn, the $h(x)$ associated with the new sample will be used to draw the next sample.
Reversible-jump Markov chain Monte Carlo

In synthesis:

There is a different $h(x)$ associated to each possible state

After a new sample is drawn, the $h(x)$ associated with the new sample will be used to draw the next sample

This sample dependent $h(x)$ is what we called transition probability:

$$\Pr(X_{n+1} = x | X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \Pr(X_{n+1} = x | X_n = x_n).$$
Reversible-jump Markov chain Monte Carlo

Where is the Markov Directed Graph? :
Reversible-jump Markov chain Monte Carlo

Where is the Markov Directed Graph? :

The Graph States set is the sampling space
Reversible-jump Markov chain Monte Carlo

Where is the Markov Directed Graph? :

The Graph States set is the sampling space

The Edges Set is given by each possible transition in the sampling space
Reversible-jump Markov chain Monte Carlo

Where is the Markov Directed Graph? :

The Graph States set is the sampling space

The Edges Set is given by each possible transition in the sampling space

Ex: continuous sampling space:
   Infinite nodes
   Infinite outgoing edges from each state
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

MCMC Specific instance
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

MCMC Specific instance

Nicholas Metropolis
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

MCMC Specific instance

Nicholas Metropolis

Colleague at the Manhattan project

Stanislaw Ulam
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

MCMC Specific instance

While Ulam's uncle was gambling his money at the Monte Carlo Casino

Nicholas Metropolis

Stanislaw Ulam

Colleague at the Manhattan project
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

Need to choose a proposal distribution $P(x|x_{k-1})$ to create the Markov chain.
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

Need to choose a proposal distribution $P(x|x_{k-1})$ to create the Markov chain

Draw sample $x^*$ from $P(x|x_{k-1})$
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

Need to choose a proposal distribution $P(x|x_{k-1})$ to create the Markov chain.

Draw sample $x^*$ from $P(x|x_{k-1})$.

Accept $x^*$ with probability:

$$\alpha(x^*, x_{k-1}) = \min\{1, \frac{f(x^*)P(x_{i-1}|x^*)}{f(x_{i-1})P(x^*|x_{i-1})}\}$$
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

As in Acceptance Rejection Sampling we need to draw from \( \text{uniform}(0, 1) \) to decide if accept
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

As in Acceptance Rejection Sampling we need to draw from \( \text{uniform}(0, 1) \) to decide if accept.

Unlike Acceptance Rejection Sampling, rejected samples are used as \( x_{k-1} \) for next sample.
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ \alpha(x^*, x_{i-1}) = \min\{1, \frac{f(x^*)P(x_{i-1}|x^*)}{f(x_{i-1})P(x^*|x_{i-1})} \} \]

What does it mean?
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[
\alpha(x^*, x_{k-1}) = \min\{1, \frac{f(x^*) P(x_{i-1} | x^*)}{f(x_{i-1}) P(x^* | x_{i-1})}\}
\]

Increase probability of acceptance if new sample \(x^*\) has higher probability than previous sample \(x_{i-1}\)
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ \alpha(x^*, x_{k-1}) = \min\{1, \frac{f(x^*) P(x_{i-1} | x^*)}{f(x_{i-1}) P(x^* | x_{i-1})} \} \]

- Increase probability of acceptance if new sample \( x^* \) has higher probability than previous sample \( x_{i-1} \)

- Increase probability of acceptance if transition from \( x^* \) to \( x_{i-1} \) is more probable than transition from \( x_{i-1} \) to \( x^* \)
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ P(x|x_{k-1}) \] model design choice to fit the problem
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ P(x | x_{k-1}) \] model design choice to fit the problem options:
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ P(x \mid x_{k-1}) \] model design choice to fit the problem options:

1) \( P(x \mid x_{k-1}) \sim N(x_{k-1}, \sigma^2 I) \)  
   Gaussian distribution centered \( x_{k-1} \)
   Is symmetric \( P(x_{i-1} \mid x^*) = P(x^* \mid x_{i-1}) \)
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ P(x|x_{k-1}) \] model design choice to fit the problem options:

1) \( P(x|x_{k-1}) \sim N(x_{k-1}, \sigma^2 I) \) Gaussian distribution centered \( x_{k-1} \)

Is symmetric \( P(x_{i-1}|x^*) = P(x^*|x_{i-1}) \)

2) \( P(x|x_{k-1}) = P(x) \) Does not depend on \( x_{i-1} \)

Called "Independence Chain Metropolis-Hastings", asymmetric
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

\[ P(x \mid x_{k-1}) \] model design choice to fit the problem options:

1) \[ P(x \mid x_{k-1}) \sim N(x_{k-1}, \sigma^2 I) \] Gaussian distribution centered \( x_{k-1} \)
   Is symmetric \( P(x_{i-1} \mid x^*) = P(x^* \mid x_{i-1}) \)

2) \[ P(x \mid x_{k-1}) = P(x) \] Does not depend on \( x_{i-1} \)
   Called "Independence Chain Metropolis-Hastings", asymmetric

3) \[ P(x \mid x_{k-1}) \] Random walk chain
   Called "Random walk Metropolis-Hastings"
Reversible-jump Markov chain Monte Carlo

**Metropolis-Hastings Sampling**

Independence walk chain

Random walk chain
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

Independence walk chain
Preferred if it is easy to produce new samples
Cover search space faster

Random walk chain
Preferred if it is easier to generate new samples as transformation of previous
Ex: parse tree or high dimensions
Reversible-jump Markov chain Monte Carlo

Metropolis-Hastings Sampling

Random walk chain adaptation for argmax solution
Referred to as "Random Search Optimization"
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling
Random Walk Metropolis-Hastings instance
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling
Random Walk Metropolis-Hastings instance

Josiah Willard Gibbs
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling
Random Walk Metropolis-Hastings instance

Josiah Willard Gibbs

Lived ~70 years before Ulam's uncle started gambling in Monte Carlo

Spare some CHANGE??
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling

Extremely useful for sampling from high dimensional distributions
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling

Extremely useful for sampling from high dimensional distributions

Change one dimension at the time
Proposal distribution conditions on all other dimensions
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling

Extremely useful for sampling from high dimensional distributions

Change one dimension at the time
Proposal distribution conditions on all other dimensions

Samples are stored after all dimensions have been changed
Gibbs Sampling

Extremely useful for sampling from high dimensional distributions

Change one dimension at the time
Proposal distribution conditions on all other dimensions

Samples are stored after all dimensions have been changed

Acceptance rate is always $= 1$
No rejections
Collect samples faster
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling

\( f(\mathbf{x}) \) is the PDF of \( \mathbf{x}, \quad \mathbf{x} = (x^1, x^2, \ldots, x^m)' \) is \( m \)-dimensional

- for \( i = 1 : n \)
  - for \( j = 1 : m \)
    * draw \( x^j_i \) from \( f(x_j|x^1_i, x^2_i, \ldots, x^{j-1}_i, x^{j+1}_i, \ldots, x^m_{i-1}) \)
  - end for

- end for

\( (x_1, x_2, \ldots, x_n) \) are samples from \( f(\mathbf{x}) \)
Reversible-jump Markov chain Monte Carlo

**Gibbs Sampling**

Sampling with Gibbs sampler
Reversible-jump Markov chain Monte Carlo

Gibbs Sampling

Gibbs sampler adapted for argmax (hill climbing)
Referred to as "Univariate Search Optimization"
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo
Reversible-jump Markov chain Monte Carlo

MCMC can be used for Bayesian modeling

Predictions require marginalization

\[ p(y|X) = \int p(y|\theta)p(\theta|X) \, d\theta \]

Marginal likelihood

Normalization factor
MCMC for Bayesian computation are restricted to problems where the number of parameters is fixed (fixed number of dimensions)
Reversible-jump Markov chain Monte Carlo

MCMC for Bayesian computation are restricted to problems where the number of parameters is fixed (fixed number of dimensions)

Not been available for application to Bayesian model selection where the number of parameters have to be chosen
Reversible-jump Markov chain Monte Carlo

MCMC for Bayesian computation are restricted to problems where the number of parameters is fixed (fixed number of dimensions)

Not been available for application to Bayesian model selection where the number of parameters have to be chosen

RJMCMC sampler can jump between parameter subspaces of differing dimensionality
Reversible-jump Markov chain Monte Carlo

Example: Change-point Analysis

A change-point model allows different parts of a dataset to obey different probability laws.

For example:

All samples drawn during time interval X follow model A,
While samples drawn during time interval Y follow model B.
Reversible-jump Markov chain Monte Carlo

Example: Change-point Analysis

A change-point model allows different parts of a dataset to obey different probability laws.

Constant intervals + noise, example:
- The target function is constant in time intervals
- Samples value carry some noise.
Reversible-jump Markov chain Monte Carlo

Example: Change-point Analysis

A change-point model allows different parts of a dataset to obey different probability laws

Need to find the most probable Change-point Model!
Reversible-jump Markov chain Monte Carlo

Example: Change-point Analysis

A change-point model allows different parts of a dataset to obey different probability laws.

Need to find:
- Number of change-points: $k$
- Position of change-points: $|k|$
- Value for each interval: $|k + 1|$

Size not fixed!

\[ \theta^{(k)} \]

parameters

observed data $y$
Reversible-jump Markov chain Monte Carlo

Example: Change-point Analysis

A change-point model allows different parts of a dataset to obey different probability laws

Solution?

If likelihood is the negative squared-error
Max likelihood always prefer more change-points
Over-fitting the data
Reversible-jump Markov chain Monte Carlo

Example: Change-point Analysis

A change-point model allows different parts of a dataset to obey different probability laws.

Solution!

Find the model that maximize: \( p(k, \theta^{(k)} | y) = p(k | y)p(\theta^{(k)} | k, y) \)

Use sampler based search.
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

Current state/model is $\mathbf{x} = (k, \theta^{(k)})$
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

Current state/model is \( x = (k, \theta^{(k)}) \)

Sampler can propose jump to state/model \( x' \) of new type \( k = m \)
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

Current state/model is $x = (k, \theta^{(k)})$

Sampler can propose jump to state/model $x'$ of new type $k = m$

The jump is accepted with probability $\alpha_m(x, x')$
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

\[ \alpha_m(x, x') \] must be chosen to attain "detailed balance" for each move type.
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

\[ \alpha_m(x, x') \]

must be chosen to attain "detailed balance" for each move type

Sampler must define transformation of parameter vector to new size
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

\[ \alpha_m(x, x') \]

must be chosen to attain "detailed balance" for each move type.

Sampler must define transformation of parameter vector to new size.

Example:

space \[ \{2\} \times \mathbb{R}^2 \rightarrow \{1\} \times \mathbb{R} \]

state \[ 2, \theta_1, \theta_2 \rightarrow 1, \frac{1}{2}(\theta_1 + \theta_2) \]
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

Adaptation of Metropolis-Hastings sampler

So that it can jump between models with different dimensionalities
Reversible-jump Markov chain Monte Carlo

Reversible-Jump sampler:

Adaptation of Metropolis-Hastings sampler

So that it can jump between models with different dimensionalities
Reversible-jump Markov chain Monte Carlo