Markov Chain Monte Carlo in Statistics — Recent Advances

Junni Zhang
Assistant Professor
Guanghua School of Management
Peking University

Nov. 3, 2003

Outline

- Design Principle of a MCMC Scheme
- Generalized Gibbs
- Multipoint Metropolis Methods
- Evolutionary Monte Carlo
- Perfect Sampling
- Dynamic Weighting

One Basic Problem of Monte Carlo

- Draw random variable

$$\boldsymbol{x} \sim \pi(\boldsymbol{x})$$

often used for simulating complex systems

- It may not be easy to draw ***x*** directly!

⟹ MCMC (Markov Chain Monte Carlo).

A MCMC Scheme

- A MCMC sampler with transition functions $A_i(\boldsymbol{x},\boldsymbol{y})$

$$\boldsymbol{x}^{(0)} \xrightarrow{\boldsymbol{x}^{(1)} \sim A_1(\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)})} \boldsymbol{x}^{(1)} \xrightarrow{\boldsymbol{x}^{(2)} \sim A_2(\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)})} \boldsymbol{x}^{(2)} \cdots$$

- Key Theory:

  If a Markov chain
  - irreducible
  - aperiodic
  - possesses an invariant distribution π

  then the chain will become stationary at π.

- Principle: Design a transition function $A(\boldsymbol{x}, \boldsymbol{y})$ that leaves the target distribution $\pi(\boldsymbol{x})$ invariant.

Invariance and Detailed Balance

- $A(\mathbf{x}, \mathbf{y})$ leaves $\pi(\mathbf{x})$ invariant if

$$\int \pi(\mathbf{x}) A(\mathbf{x}, \mathbf{y}) d\mathbf{x} = \pi(\mathbf{y})$$

$$\mathbf{x}^{(t)} = \mathbf{x} \sim \pi \xrightarrow{y \sim A(\mathbf{x}, y)} \mathbf{x}^{(t+1)} = \mathbf{y} \sim \pi$$

- Detailed balance:

$$\pi(\mathbf{x}) A(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{y}) A(\mathbf{y}, \mathbf{x})$$

It ensures invariance since

$$\int \pi(\mathbf{x}) A(\mathbf{x}, \mathbf{y}) d\mathbf{x} = \int \pi(\mathbf{y}) A(\mathbf{y}, \mathbf{x}) d\mathbf{x} = \pi(\mathbf{y}) \int A(\mathbf{y}, \mathbf{x}) d\mathbf{x} = \pi(\mathbf{y})$$

Gibbs Sampler

- Purpose: Draw from a Joint Distribution

$$\boldsymbol{x} = (x_1, \cdots, x_d) \sim \pi(\boldsymbol{x})$$

- Method: Iterative Conditional Sampling

$$\forall i \quad \boldsymbol{x} = (x_i, \boldsymbol{x}_{[-i]})$$

$$\text{draw} \quad x_i' \sim \pi(x_i' \mid \boldsymbol{x}_{[-i]})$$

$$\text{let} \quad \boldsymbol{y} = (x_i', \boldsymbol{x}_{[-i]})$$

Metropolis-Hastings Algorithm

- Draw $\boldsymbol{y}$ from a proposal distribution $T(\boldsymbol{x},\boldsymbol{y})$.

- Accept $\boldsymbol{y}$ with probability

$$r(\boldsymbol{x},\boldsymbol{y}) = \min\left\{1, \frac{\pi(\boldsymbol{y})T(\boldsymbol{y},\boldsymbol{x})}{\pi(\boldsymbol{x})T(\boldsymbol{x},\boldsymbol{y})}\right\}$$

  and stay at $\boldsymbol{x}$ with probability $1-r(\boldsymbol{x},\boldsymbol{y})$.

Note: One can check that detailed balance is satisfied.

Beyond Invariance

Make the Markov chain explore the relevant space quickly and reach stationarity quickly!

Outline

- Design Principle of a MCMC Scheme
- Generalized Gibbs
- Multipoint Metropolis Methods
- Evolutionary Monte Carlo
- Perfect Sampling
- Dynamic Weighting

Gibbs Sampler

$\forall i \quad \boldsymbol{x} = (x_i, \boldsymbol{x}_{[-i]})$

draw $x_i' \sim \pi(x_i' \mid \boldsymbol{x}_{[-i]})$

let $\boldsymbol{y} = (x_i', \boldsymbol{x}_{[-i]})$

⟺

draw $\gamma \sim \pi(x_i + \gamma \mid \boldsymbol{x}_{[-i]})$

let $\boldsymbol{y} = \gamma(\boldsymbol{x}) = (x_i + \gamma, \boldsymbol{x}_{[-i]})$

a transformation

Group Moves

Try moving along an arbitrary direction…

Given any fixed direction $\boldsymbol{e} = (e_1, \cdots, e_d)$

$$\gamma(\boldsymbol{x}) = \boldsymbol{x} + \gamma\boldsymbol{e} = (x_1 + \gamma e_1, \cdots, x_d + \gamma e_d)$$

Try scaling…

$$\gamma(\boldsymbol{x}) = \gamma\boldsymbol{x} = (\gamma x_1, \cdots, \gamma x_d)$$

* Generalized Gibbs

Form of the transition function $A(\boldsymbol{x},\boldsymbol{y})$:

- selecting a transformation $\gamma \in \Gamma$
- letting $\boldsymbol{y} = \gamma(\boldsymbol{x})$.

Question: What distribution should one draw $\gamma \in \Gamma$ from so that $\pi(\boldsymbol{x})$ is left invariant?

A Theorem

- Γ={all γ} is a locally compact group
- L is its left-Haar measure
- $\gamma \sim p_{\boldsymbol{x}}(\gamma) \propto \pi(\gamma(\boldsymbol{x}))\,|J_\gamma(\boldsymbol{x})|\,L(d\gamma)$

  where

$$J_\gamma(\boldsymbol{x}) = \det\left\{\frac{\partial \gamma(\boldsymbol{x})}{\partial \boldsymbol{x}}\right\}$$

  is the Jacobian of the transformation

If $\boldsymbol{x} \sim \pi(\boldsymbol{x})$, then $\boldsymbol{y} = \gamma(\boldsymbol{x}) \sim \pi$.

Note: A left-Haar measure satisfies

$$\forall \gamma_0 \in \Gamma, \forall B \subset \Gamma, L(B) = L(\gamma_0 B)$$

Translation group along an arbitrary direction…

Given any fixed direction $\boldsymbol{e} = (e_1, \cdots, e_d)$

$$\Gamma = \left\{ \gamma \in R^1 : \gamma(\boldsymbol{x}) = \boldsymbol{x} + \gamma \boldsymbol{e} = (x_1 + \gamma e_1, \cdots, x_d + \gamma e_d) \right\}$$

Draw $\gamma \sim p_{\boldsymbol{x}}(\gamma) \propto \pi(\boldsymbol{x} + \gamma \boldsymbol{e})$

Let $\boldsymbol{y} = \gamma(\boldsymbol{x}) = \boldsymbol{x} + \gamma \boldsymbol{e}$

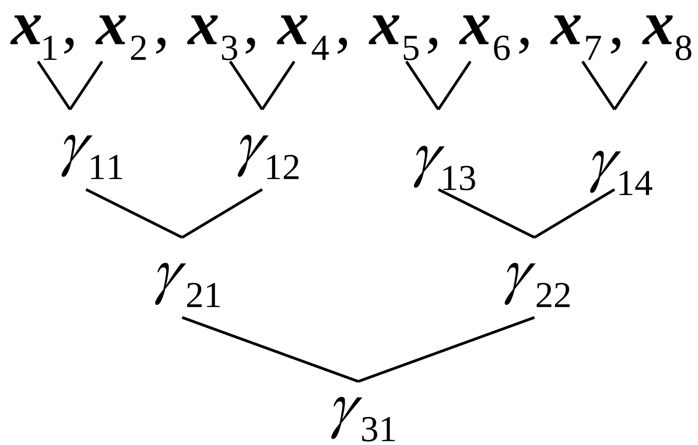
Scale-transformation Group…

$$\Gamma = \left\{ \gamma \in R^1 \setminus \{0\} : \gamma(\boldsymbol{x}) = (\gamma x_1, \cdots, \gamma x_d) \right\}$$

Draw $\gamma \sim p_{\boldsymbol{x}}(\gamma) \propto |\gamma|^{d-1} \pi(\gamma \boldsymbol{x})$

Let $\boldsymbol{y} = \gamma(\boldsymbol{x}) = \gamma \boldsymbol{x}$

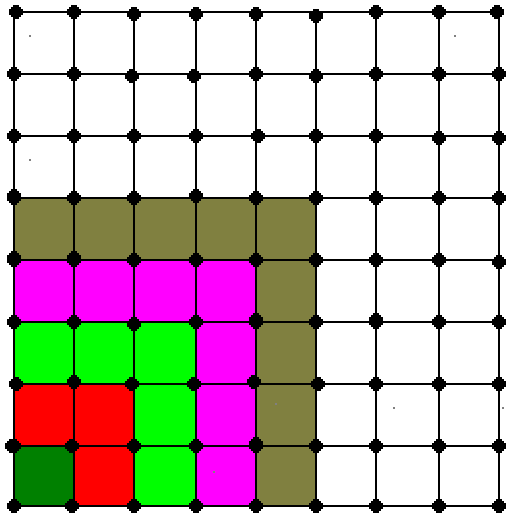
More Examples of Group Moves



In general, for any subset S of $\{1,\ldots,8\}$,
$\{\boldsymbol{x}_i, i \in S\}$ can be moved together.

$$z = \{z_\sigma, \text{all lattice points}\}$$

$$\pi(z) \propto e^{-\beta H(z)}$$



In general, for any subset S of the lattice points, $\{z_\sigma, \sigma \in S\}$ can be moved together.

If we can't directy draw from $p_{\boldsymbol{x}}(\gamma)$

- MCMC transition function $A_{\boldsymbol{x}}(\gamma, \gamma')$

—— need to leave $p_{\boldsymbol{x}}(\gamma)$ invariant

—— need to be "transformation invariant"

$$A_{\boldsymbol{x}}(\gamma, \gamma') = A_{\gamma_0(\boldsymbol{x})}(\gamma\gamma_0^{-1}, \gamma'\gamma_0^{-1})$$

group operation

Outline

- Design Principle of a MCMC Scheme
- Generalized Gibbs
- Multipoint Metropolis Methods
- Evolutionary Monte Carlo
- Perfect Sampling
- Dynamic Weighting

Difficulty of Choosing The Proposal Distribution

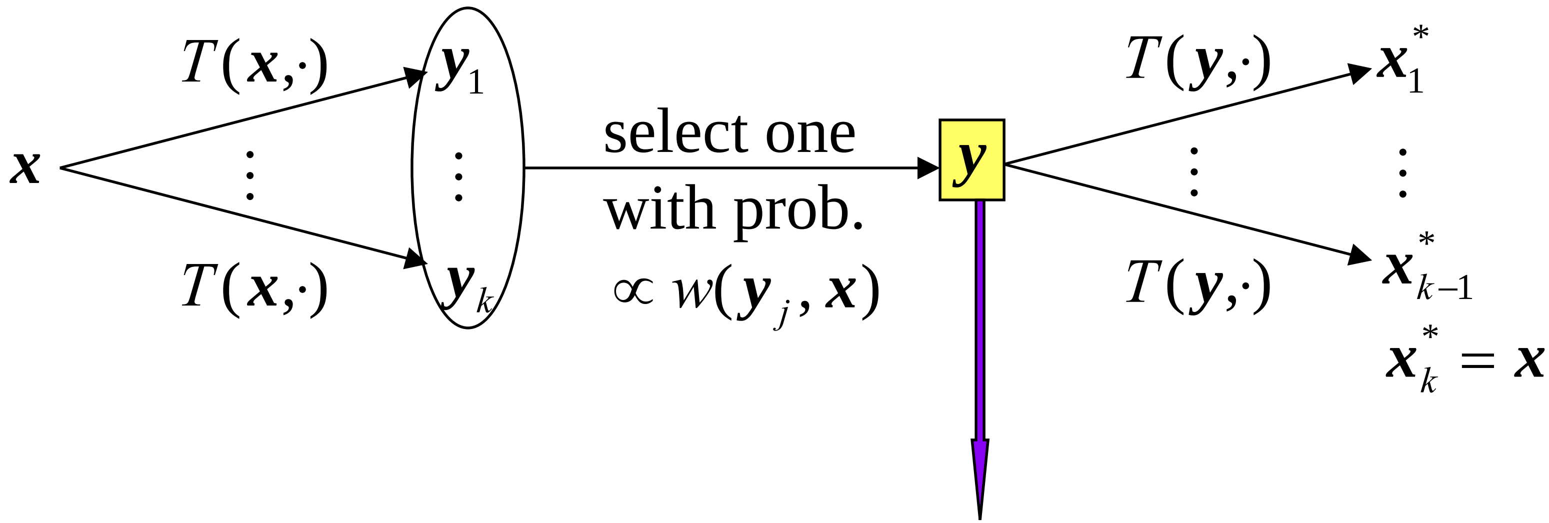
- small step size in the proposal distribution
  ⟹ slow movement of the Markov chain
- large step size in the proposal distribution
  ⟹ low acceptance rate

In both cases, the chain moves slowly!

* Multipoint Metropolis Methods

- Idea: make multiple proposals and select a good one from them.
- Guiding principle: leaving the target distribution $\pi(\boldsymbol{x})$ invariant!

Independent Multipoint Proposals



accept with prob. $r_g = \min\left\{1, \dfrac{w(\boldsymbol{y}_1, \boldsymbol{x}) + \cdots + w(\boldsymbol{y}_k, \boldsymbol{x})}{w(\boldsymbol{x}_1^*, \boldsymbol{y}) + \cdots + w(\boldsymbol{x}_k^*, \boldsymbol{y})}\right\}$

reject with prob. $1 - r_g$

- Remark 1:

$$w(\boldsymbol{x}, \boldsymbol{y}) = \pi(\boldsymbol{x}) T(\boldsymbol{x}, \boldsymbol{y}) \lambda(\boldsymbol{x}, \boldsymbol{y})$$

  where $\lambda(\boldsymbol{x}, \boldsymbol{y})$ is a non-negative symmetric function that can be chosen by the user.

- Remark 2: Detailed balance is satisfied.

Correlated Multipoint Proposals

select one with prob.

$$\propto w(\boldsymbol{y}_{[l:1]}, \boldsymbol{x})$$

suppose $\boldsymbol{y} = \boldsymbol{y}_j$

$$P_j(\cdot \mid \boldsymbol{x}, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_{j-1})$$

$$\boldsymbol{x} \rightarrow \boldsymbol{y}_1 \rightarrow \cdots \rightarrow \boldsymbol{y}_{j-1} \rightarrow \boldsymbol{y}_j \rightarrow \cdots \rightarrow \boldsymbol{y}_k$$

$$\boldsymbol{x}_k^* \leftarrow \cdots \leftarrow \boldsymbol{x}_m^* \leftarrow \cdots \leftarrow \boldsymbol{x}_{j+1}^* \leftarrow \boldsymbol{x}_j^* \leftarrow \boldsymbol{x}_{j-1}^* \leftarrow \cdots \leftarrow \boldsymbol{x}_1^* \leftarrow \boldsymbol{y}$$

$$P_m(\cdot \mid \boldsymbol{y}, \boldsymbol{x}_1^*, \cdots, \boldsymbol{x}_{m-1}^*)$$

accept with prob. $r = \min\left\{1, \dfrac{w(\boldsymbol{y}_1, \boldsymbol{x}) + \cdots + w(\boldsymbol{y}_{[k:1]}, \boldsymbol{x})}{w(\boldsymbol{x}_1^*, \boldsymbol{y}) + \cdots + w(\boldsymbol{x}_{[k:1]}^*, \boldsymbol{y})}\right\}$

reject with prob. $1 - r$

- Remark 1:

$$P_j(\boldsymbol{y}_{[j:1]} \mid \boldsymbol{x}) = P_1(\boldsymbol{y}_1 \mid \boldsymbol{x}) \cdots P_j(\boldsymbol{y}_j \mid \boldsymbol{x}, \boldsymbol{y}_{[1:j-1]})$$

$$w_j(\boldsymbol{y}_{[j:1]}, \boldsymbol{x}) = \pi(\boldsymbol{x}) P_j(\boldsymbol{y}_{[j:1]} \mid \boldsymbol{x}) \lambda_j(\boldsymbol{x}, \boldsymbol{y}_{[1:j]})$$

  where $\lambda_j(\boldsymbol{x}, \boldsymbol{y}_{[1:j]})$ is a sequentially symmetric function that can be chosen by the user:

$$\lambda_j(a, b, \cdots, z) = \lambda_j(z, \cdots, b, a)$$

- Remark 2: Detailed balance is satisfied.

Outline

- Design Principle of a MCMC Scheme
- Generalized Gibbs
- Multipoint Metropolis Methods
- Evolutionary Monte Carlo
- Perfect Sampling
- Dynamic Weighting

Population-based "Learning" Strategy

- conduct parallel Monte Carlo Markov chains
- interactions among the multiple chains in the "population"
  - mutation
  - crossover
  - exchange

  purpose: improve "fitness" of the members

* Evolutionary Monte Carlo In A Tempering Framework

- Target distribution:

$$\pi(\boldsymbol{x}) \propto \exp\{-H(\boldsymbol{x})\}$$

- Population: $\boldsymbol{X}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_m\}$

$$\pi(\boldsymbol{x}_i) \propto \exp\{-H(\boldsymbol{x}_i)/t_i\}$$

$1=t_1<t_2<\ldots<t_m$

- Target distribution of the population: the augmented Boltzmann distribution

$$\pi(\boldsymbol{X}) \propto \exp\{-\sum_{i=1}^{m} H(\boldsymbol{x}_i)/t_i\}$$

- Example: binary-coded state space

$$\boldsymbol{x}_i = (b_{i,1},\cdots,b_{i,d}), \quad i=1,\cdots,m$$

Mutation

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_m \end{pmatrix} \xrightarrow[\text{selection}]{\text{random}} \boldsymbol{x}_i \longrightarrow$$

1-point mutation: randomly select j

$$\boldsymbol{x}_i = (b_{i,1}, \cdots, \quad b_{i,j} \quad, \cdots, b_{i,d})$$

$$\boxed{\boldsymbol{y}_i} = (b_{i,1}, \cdots, 1 - b_{i,j}, \cdots, b_{i,d})$$

$$\boxed{\boldsymbol{Y} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{y}_i, \cdots, \boldsymbol{x}_m\}}$$

accept with prob.

$$\min\left\{1, \frac{\pi(\boldsymbol{Y})}{\pi(\boldsymbol{X})} = \exp\{-(H(\boldsymbol{y}_k) - H(\boldsymbol{x}_k))/t_k\right\}$$

Crossover

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_m \end{pmatrix} \longrightarrow \text{select } \boldsymbol{x}_i, \boldsymbol{x}_j \text{ with}$$

$$P((\boldsymbol{x}_i, \boldsymbol{x}_j) \mid \boldsymbol{X}) \propto [\exp\{-H(\boldsymbol{x}_i)/t_s\} + \exp\{-H(\boldsymbol{x}_j)/t_s\}] \quad \boldsymbol{x}_i \neq \boldsymbol{x}_j$$

→ 1-point crossover: randomly select position k

$$\boldsymbol{x}_i = (b_{i,1}, \cdots, b_{i,k}, \cdots, b_{i,d}) \quad \boldsymbol{x}_j = (b_{j,1}, \cdots, b_{j,k}, \cdots, b_{j,d})$$

$$\boldsymbol{y}_i = (b_{i,1}, \cdots, b_{j,k}, \cdots, b_{i,d}) \quad \boldsymbol{y}_j = (b_{j,1}, \cdots, b_{i,k}, \cdots, b_{j,d})$$

$$\boldsymbol{Y} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{y}_i, \cdots, \boldsymbol{y}_j, \cdots, \boldsymbol{x}_m\}$$

accept with prob.

$$\min\left\{1, \frac{\pi(\boldsymbol{Y})T(\boldsymbol{Y},\boldsymbol{X})}{\pi(\boldsymbol{X})T(\boldsymbol{X},\boldsymbol{Y})} = \exp\left\{-\frac{H(\boldsymbol{y}_i)-H(\boldsymbol{x}_i)}{t_i} - \frac{H(\boldsymbol{y}_j)-H(\boldsymbol{x}_j)}{t_j}\right\}\frac{T(\boldsymbol{Y},\boldsymbol{X})}{T(\boldsymbol{X},\boldsymbol{Y})}\right\}$$

Exchange

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_m \end{pmatrix} \xrightarrow[\text{selection}]{\text{random}} \boldsymbol{x}_i, \boldsymbol{x}_j$$

$\longrightarrow$ exchange

$$\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_j, \cdots, \boldsymbol{x}_m\}$$

$$\boldsymbol{Y} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_j, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_m\}$$

accept with prob.

$$\min\left\{1, \frac{\pi(\boldsymbol{Y})}{\pi(\boldsymbol{X})} = \exp\{-(H(\boldsymbol{x}_i) - H(\boldsymbol{x}_j))(1/t_j - 1/t_i)\}\right\}$$

Continuous Sample Space: Mutation

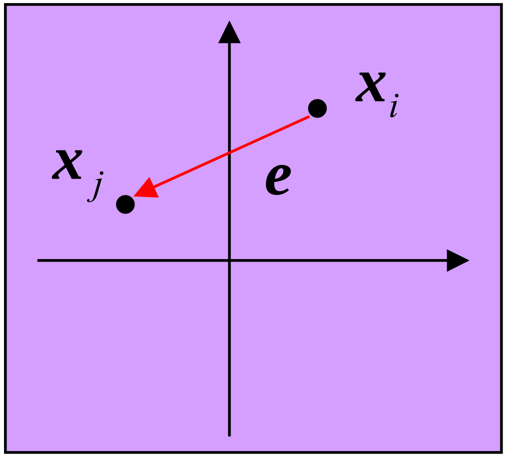
Any kind of Metropolis-Hasting move independently for each chain!

Continuous Sample Space: Snooker Crossover

$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_m \end{pmatrix}$ —random selection→ $\boldsymbol{x}_i$ → select "anchor"

$\boldsymbol{x}_j \in \boldsymbol{X} \setminus \{\boldsymbol{x}_i\}$ with prob.

$\propto \exp\{-H(\boldsymbol{x}_j)/t_s\}$



→ let $\boldsymbol{e} = \dfrac{\boldsymbol{x}_j - \boldsymbol{x}_i}{\|\boldsymbol{x}_j - \boldsymbol{x}_i\|}$

draw $\gamma \sim p(\gamma) \propto |\gamma|^{d-1}\, \pi(\boldsymbol{x}_i + \gamma \boldsymbol{e})$

let $\boldsymbol{y}_i = \pi(\boldsymbol{x}_i + \gamma \boldsymbol{e})$

new population

$\boldsymbol{Y} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{y}_i, \cdots, \boldsymbol{x}_m\}$

Outline

- Design Principle of a MCMC Scheme
- Generalized Gibbs
- Multipoint Metropolis Methods
- Evolutionary Monte Carlo
- Perfect Sampling
- Dynamic Weighting

An Observation

- If a Markov chain had been started from the infinite past

starting point

$t = -\infty$ …… $t = 0$

$\boldsymbol{x}_{-\infty}$ $\boldsymbol{x}_0 \sim \pi$

stationary

The Idea of Coupling

- Assume finite Markov chain: $\chi = \{1, \cdots, |\chi|\}$

$t = -1$ $\qquad$ $t = 0$

$\boldsymbol{x}_{-1}$ $\qquad$ $\boldsymbol{x}_0 \sim A(\boldsymbol{x}_{-1}, \cdot)$ $\qquad$ $\boldsymbol{x}_0 = \phi(u_0, \boldsymbol{x}_{-1})$

(1) Compute

$$G(\boldsymbol{x}_{-1}, j) = \sum_{k=1}^{j} A(\boldsymbol{x}_{-1}, j) = \Pr(\boldsymbol{x}_0 \leq j \mid \boldsymbol{x}_{-1})$$

(2) generate u_0~*Uniform(0,1)*

(3) Let $\boldsymbol{x}_0 = j$ if

$$G(\boldsymbol{x}_{-1}, j-1) < u_0 \leq G(\boldsymbol{x}_{-1}, j)$$

- The chains starting from all possible states are *t=-1* are ***coupled by the same random number*** u_0.

* Perfect Sampling

- If u_0 makes all the chain "coupled", that is,

$$\forall i \quad \phi(u_0, i) \equiv j_0$$

then $\boldsymbol{x}_0 = j_0 \sim \pi$

starting point

$t = -\infty$	……	$t = -1$	$t = 0$	stationary
$\boldsymbol{x}_{-\infty}$		$\exists i \quad \boldsymbol{x}_{-1} = i$	$\boldsymbol{x}_0 \equiv j_0$	$\sim \pi$

If the chains are not coupled in one step…

$$\boldsymbol{x}_{-(n-1)} = \phi(u_{-(n-1)}, \boldsymbol{x}_{-n})$$

$$\Rightarrow \boldsymbol{x}_0 = \phi(u_0, \phi(u_{-1}, \cdots, \phi(u_{-(n-1)}, \boldsymbol{x}_{-n})\cdots))$$

- Equivalently,
  - The sequence of uniform random variables

    $$\cdots, u_{-n}, \cdots, u_{-1}, u_0$$

    are given in advance
  - From the infinite past, we compose

    $$\cdots, \boldsymbol{x}_{-n}, \cdots, \boldsymbol{x}_{-1}, \boldsymbol{x}_0$$

- If $\forall i \quad \phi(u_0, \phi(u_{-1}, \cdots, \phi(u_{-(n-1)}, i)\cdots)) = j_0$

  then $\boldsymbol{x}_0 = j_0 \sim \pi$

starting point

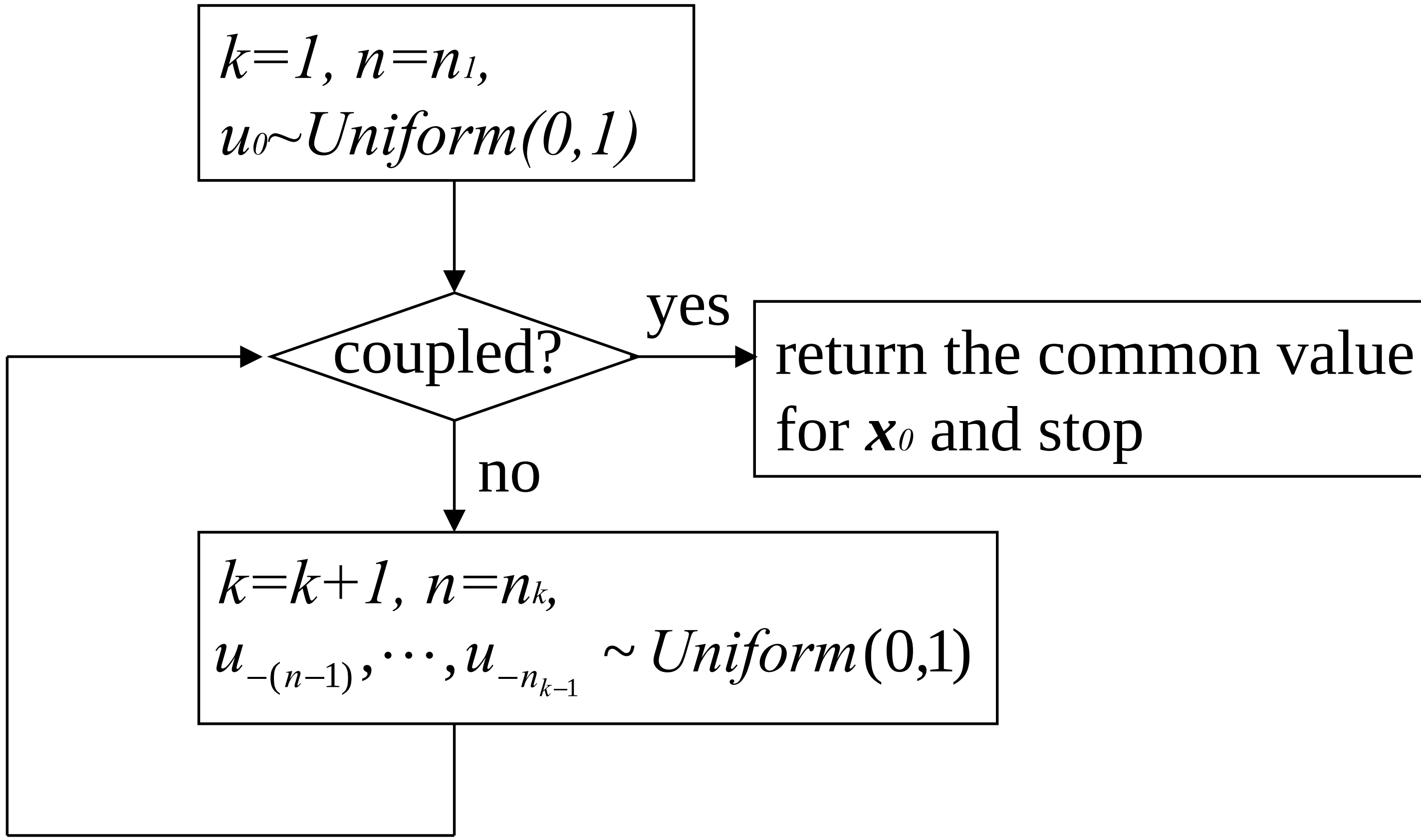
$t = -\infty$ …… $t = -n$ …… $t = 0$

$\boldsymbol{x}_{-\infty}$ $\quad \exists i \quad \boldsymbol{x}_{-n} = i$ $\quad \boldsymbol{x}_0 \equiv j_0$ $\sim \pi$

stationary

The Conceptual Algorithm

$$\text{schedule}: \ 1 = n_1 < n_2 < n_3 < \cdots$$



$$\cdots\cdots \quad u_{-(n_k-1)}, \cdots, u_{-n_{k-1}} \quad \cdots\cdots \quad u_{-(n_2-1)}, \cdots, u_{-1} \quad u_0$$

Outline

- Design Principle of a MCMC Scheme
- Generalized Gibbs
- Multipoint Metropolis Methods
- Evolutionary Monte Carlo
- Perfect Sampling
- Dynamic Weighting

Weighted Sample

- Augment the sample space from $\mathcal{X}$ to $\chi \times R^+$ to include a weight variable for each state.

$$\boldsymbol{x} \rightarrow (\boldsymbol{x}, w)$$

- Suppose $f(\boldsymbol{x},w)$ is the joint distribution of $(\boldsymbol{x},w)$. $\boldsymbol{x}$ is correctly weighted by w with respect to π if

$$\sum_w wf(\boldsymbol{x}, w) \propto \pi(\boldsymbol{x})$$

Note: we can estimate $E_\pi[h(\boldsymbol{x})]$ by

$$\frac{\sum_{i=1}^n w_i h(\boldsymbol{x}_i)}{\sum_{i=1}^n w_i} \quad \text{where } (\boldsymbol{x}_i, w_i) \sim f, i = 1, \cdots, n$$

New Design Principle

- IWIW (Invariance With respect to Importance Weighting).
  - A transition rule $A(\boldsymbol{x},w;\boldsymbol{y},w')$ satisfies IWIW if

$$f(\boldsymbol{x},w) \xrightarrow{A(\boldsymbol{x},w;\boldsymbol{y},w')} f'(\boldsymbol{y},w')$$

correctly weighted wrt π

correctly weighted wrt π

* Dynamic Weighting

M-type move: Given $(\boldsymbol{x}_t, w_t) = (\boldsymbol{x}, w)$ at iteration t.

- Draw $\boldsymbol{x}_{t+1}$ from a transition function that leaves π invariant.

- Set $w_{t+1}=w$.

R-type move: Given $(\boldsymbol{x}_t, w_t) = (\boldsymbol{x}, w)$ at iteration t.

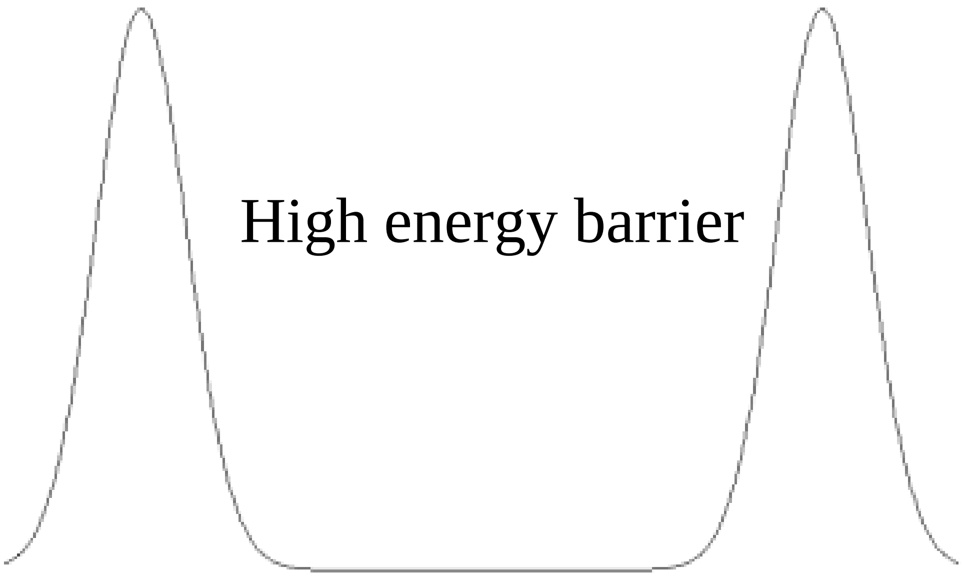
- Propose $\boldsymbol{y}$ from $T(\boldsymbol{x},\boldsymbol{y})$ and compute

$$r(\boldsymbol{x},\boldsymbol{y}) = \frac{\pi(\boldsymbol{y})T(\boldsymbol{y},\boldsymbol{x})}{\pi(\boldsymbol{x})T(\boldsymbol{x},\boldsymbol{y})}$$

- Choose $\theta=\theta(\boldsymbol{x},w)>0$, and draw $U\sim Uniform(0,1)$. Let

$$(\boldsymbol{x}_{t+1}, w_{t+1}) = \begin{cases} (\boldsymbol{y}, wr(\boldsymbol{x},\boldsymbol{y})+\theta) & \text{if } U \le \dfrac{wr(\boldsymbol{x},\boldsymbol{y})}{wr(\boldsymbol{x},\boldsymbol{y})+\theta} \\ \left(\boldsymbol{x}, \dfrac{w(wr(\boldsymbol{x},\boldsymbol{y})+\theta)}{\theta}\right) & \text{otherwise} \end{cases}$$

Waiting Time Infinity → Importance Weight Infinity



- In the standard Metropolis process, the waiting time to cross over the barriers is infinite.
- The dynamic weighting process can cross the energy barrier, but has “importance weight infinity”.

Combinatory Strategy

- Use the weighted moves when proposing large changes in the system.
- Use the standard Metropolis or Gibbs moves for local exploration.

Reference

- Liu, J. S. (2001). Monte Carlo strategies in scientific computing. Springer-Verlag New York.
- Generalized Gibbs:
  - Liu, J. S., Sabatti, C. (2000). Generalized Gibbs sampler and multigrid Monte Carlo for Bayesian computation, *Biometrika*, 87 (2), 353-369.
- Multipoint Metropolis Methods:
  - Liu, J. S., Liang, F. and Wong, W. H. (2000). The use of multiple-try method and local optimization in Metropolis sampling, *Journal of the American Statistical Association*, 95, 121-134.
  - Qin, Z. and Liu, J. S. (2000). Multipoint Metropolis method with application to hybrid Monte Carlo, *Technical report*, Harvard University.

- Evolutionary Monte Carlo:
  - Liang, F. and Wong, W. H. (2000). Evolutionary Monte Carlo: applications to *Cp* model sampling and change point problem, *Statistica Sinica*, 10 (2), 317-342.
- Perfect Sampling:
  - Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics, *Random Structures and Algorithms*, 9 (1-2): 223-252.
- Dynamic Weighting:
  - Wong, W. H. and Liang, F. M. (1997). Dynamic weighting in Monte Carlo and optimization, *Proceedings of the National Academy of Sciences of USA*, 94 (26), 14220-14224.

Thanks!