Big Data Analysis
(MA60306)

Bibhas Adhikari

Spring 2022-23, IIT Kharagpur

Lecture 24
March 30, 2023

Random walks

Transition matrix A random process $\{X_0, X_1, \ldots\}$ with finite state space $S = \{s_1, s_2, \ldots, s_k\}$ is a (homogeneous) Markov chain with transition matrix $P = [p_{ij}] \in \mathbb{R}^{k\times k}$ if

$$p(X_{t+1} = s_j | X_t = s_i) = p_{ij}$$

for all t, $i, j \in \{1, 2, \ldots, k\}$

Random walks

Transition matrix A random process $\{X_0, X_1, \ldots\}$ with finite state space $S = \{s_1, s_2, \ldots, s_k\}$ is a (homogeneous) Markov chain with transition matrix $P = [p_{ij}] \in \mathbb{R}^{k\times k}$ if

$$p(X_{t+1} = s_j | X_t = s_i) = p_{ij}$$

for all t, $i, j \in \{1, 2, \ldots, k\}$

Example:

$$P = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

in the random walk example

Random walks

Transition matrix A random process $\{X_0, X_1, \ldots\}$ with finite state space $S = \{s_1, s_2, \ldots, s_k\}$ is a (homogeneous) Markov chain with transition matrix $P = [p_{ij}] \in \mathbb{R}^{k \times k}$ if

$$p(X_{t+1} = s_j | X_t = s_i) = p_{ij}$$

for all t, $i, j \in \{1, 2, \ldots, k\}$

Example:

$$P = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

in the random walk example

Question Does the transition matrix have any special property?

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j | X_t = i)$.

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j \,|\, X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j | X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i) p(X_1 = s_j | X_0 = s_i) = \sum_{i=1}^{k} \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j | X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^k p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^k p(X_0 = s_i) p(X_1 = s_j | X_0 = s_i) = \sum_{i=1}^k \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Question What is the long-term behavior (asymptotic) of Markov chains?

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j |\, X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i) p(X_1 = s_j |\, X_0 = s_i) = \sum_{i=1}^{k} \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Question What is the long-term behavior (asymptotic) of Markov chains?

Observation

→ Apparently, the value of X_t will keep fluctuating as $t \to \infty$

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j |\, X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i) p(X_1 = s_j |\, X_0 = s_i) = \sum_{i=1}^{k} \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Question What is the long-term behavior (asymptotic) of Markov chains?

Observation

- $\rightarrow$ Apparently, the value of X_t will keep fluctuating as $t \rightarrow \infty$
- $\rightarrow$ Can we hope that the distribution of X_t converge to a limit?

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j \mid X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^{k} p(X_0 = s_i) p(X_1 = s_j \mid X_0 = s_i) = \sum_{i=1}^{k} \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Question What is the long-term behavior (asymptotic) of Markov chains?

Observation

- $\rightarrow$ Apparently, the value of X_t will keep fluctuating as $t \rightarrow \infty$
- $\rightarrow$ Can we hope that the distribution of X_t converge to a limit?

Example Consider the third transition graph above and set $\mu^{(0)} = (\frac{1}{6}, \frac{5}{6})$. What is your observation?

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j | X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^k p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^k p(X_0 = s_i) p(X_1 = s_j | X_0 = s_i) = \sum_{i=1}^k \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Question What is the long-term behavior (asymptotic) of Markov chains?

Observation

→ Apparently, the value of X_t will keep fluctuating as $t \to \infty$

→ Can we hope that the distribution of X_t converge to a limit?

Example Consider the third transition graph above and set $\mu^{(0)} = (\frac{1}{6}, \frac{5}{6})$. What is your observation?

Question Can there exist any such other $\mu^{(0)}$? (Give it a try!)

Sampling methods

m-step transition matrix $[P^m]_{ij} = p(X_{t+m} = j \mid X_t = i)$.
Theorem $\mu^{(t)} = \mu^{(0)} P^t$
Proof(Hint) $\mu_j^{(1)} = p(X_1 = s_j) = \sum_{i=1}^k p(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^k p(X_0 = s_i) p(X_1 = s_j \mid X_0 = s_i) = \sum_{i=1}^k \mu_i^{(0)} p_{ij} = (\mu^{(0)} P)_j$

Question What is the long-term behavior (asymptotic) of Markov chains?

Observation

- $\rightarrow$ Apparently, the value of X_t will keep fluctuating as $t \rightarrow \infty$
- $\rightarrow$ Can we hope that the distribution of X_t converge to a limit?

Example Consider the third transition graph above and set $\mu^{(0)} = (\frac{1}{6}, \frac{5}{6})$. What is your observation?

Question Can there exist any such other $\mu^{(0)}$? (Give it a try!)

Sampling methods

Stationary distribution A row vector $\pi = (\pi_1, \ldots, \pi_k)$ is said to be a stationary distribution fro the Markov chain if

(i) $\pi_i \geq 0$ and $\sum_{i=1}^{k} \pi_i = 1$, and

(ii) $\pi P = \pi$, i.e. $\sum_{i=1}^{k} p_{ij} = \pi_j$, for $j = 1, \ldots, k$

Sampling methods

Stationary distribution A row vector $\pi = (\pi_1, \ldots, \pi_k)$ is said to be a stationary distribution fro the Markov chain if

(i) $\pi_i \geq 0$ and $\sum_{i=1}^{k} \pi_i = 1$, and

(ii) $\pi P = \pi$, i.e. $\sum_{i=1}^{k} p_{ij} = \pi_j$, for $j = 1, \ldots, k$

Question What is the meaning of this condition? Can we say that a Markov chain is stationary if and only if $p(X_{t+1} = y | X_t = x)$ is independent of t for any x, y?

Sampling methods

Stationary distribution A row vector $\pi = (\pi_1, \ldots, \pi_k)$ is said to be a stationary distribution fro the Markov chain if

(i) $\pi_i \geq 0$ and $\sum_{i=1}^{k} \pi_i = 1$, and

(ii) $\pi P = \pi$, i.e. $\sum_{i=1}^{k} p_{ij} = \pi_j$, for $j = 1, \ldots, k$

Question What is the meaning of this condition? Can we say that a Markov chain is stationary if and only if $p(X_{t+1} = y | X_t = x)$ is independent of t for any x, y?

Question Is every Markov chain stationary? Can there exist more than one stationary distributions of a Markov chain?

Sampling methods

reversible distribution A probability distribution π for a Markov chain on the state space $S = \{s_1, \ldots, s_k\}$ is called reversible for a Markov chain if for all $i, j \in \{1, \ldots, k\}$

$$\pi_i p_{ij} = \pi_j p_{ji}$$

Sampling methods

reversible distribution A probability distribution π for a Markov chain on the state space $S = \{s_1, \ldots, s_k\}$ is called reversible for a Markov chain if for all $i, j \in \{1, \ldots, k\}$

$$\pi_i p_{ij} = \pi_j p_{ji}$$

The Markov chain with a reversible distribution is called reversible Markov chain.

Sampling methods

reversible distribution A probability distribution π for a Markov chain on the state space $S = \{s_1, \ldots, s_k\}$ is called reversible for a Markov chain if for all $i, j \in \{1, \ldots, k\}$

$$\pi_i p_{ij} = \pi_j p_{ji}$$

The Markov chain with a reversible distribution is called reversible Markov chain.

Observation If $\pi_i p_{ij}$ is interpreted as the amount of probability mass flowing at any time instant from the state s_i to s_j, then for reversible Markov chain, it is equal to the amount of probability mass flowing from s_j to s_i. (a sense of equilibrium/balance!)

Sampling methods

reversible distribution A probability distribution π for a Markov chain on the state space $S = \{s_1, \ldots, s_k\}$ is called reversible for a Markov chain if for all $i, j \in \{1, \ldots, k\}$

$$\pi_i p_{ij} = \pi_j p_{ji}$$

The Markov chain with a reversible distribution is called reversible Markov chain.

Observation If $\pi_i p_{ij}$ is interpreted as the amount of probability mass flowing at any time instant from the state s_i to s_j, then for reversible Markov chain, it is equal to the amount of probability mass flowing from s_j to s_i. (a sense of equilibrium/balance!)
The equation $\pi_i p_{ij} = \pi_j p_{ji}$ is also called detailed balance equation.

Sampling methods

reversible distribution A probability distribution π for a Markov chain on the state space $S = \{s_1, \ldots, s_k\}$ is called reversible for a Markov chain if for all $i, j \in \{1, \ldots, k\}$

$$\pi_i p_{ij} = \pi_j p_{ji}$$

The Markov chain with a reversible distribution is called reversible Markov chain.

Observation If $\pi_i p_{ij}$ is interpreted as the amount of probability mass flowing at any time instant from the state s_i to s_j, then for reversible Markov chain, it is equal to the amount of probability mass flowing from s_j to s_i. (a sense of equilibrium/balance!)

The equation $\pi_i p_{ij} = \pi_j p_{ji}$ is also called detailed balance equation.

Theorem If π is a reversible distribution then it is also a stationary distribution

Sampling methods

Question Why do they call it ‘reversible’?

Sampling methods

Question Why do they call it 'reversible'?

Homework For a reversible MC, if the initial distribution is stationary then $P(X_t = j|\, X_{t+1} = i) = p(X_{t+1} = j|\, X_t = i)$

Sampling methods

Question Why do they call it ‘reversible’?

Homework For a reversible MC, if the initial distribution is stationary then $P(X_t = j | X_{t+1} = i) = p(X_{t+1} = j | X_t = i)$

Regular Markov chain A stationary Markov chain with transition probability matrix P is called regular (also called communicable) if there exists a $k > 0$ such that $[P^k]_{ij} > 0$ for all $i, j \in S$

Sampling methods

Question Why do they call it 'reversible'?

Homework For a reversible MC, if the initial distribution is stationary then $P(X_t = j | X_{t+1} = i) = p(X_{t+1} = j | X_t = i)$

Regular Markov chain A stationary Markov chain with transition probability matrix P is called regular (also called communicable) if there exists a $k > 0$ such that $[P^k]_{ij} > 0$ for all $i, j \in S$

Irreducibility A stationary Markov chain with transition probability matrix P is called irreducible if for any $i, j \in S$, $i \neq j$, there exists a $k > 0$ (possibly depending on i, j) such that $[P^k]_{ij} > 0$

Sampling methods

Question Why do they call it 'reversible'?

Homework For a reversible MC, if the initial distribution is stationary then $P(X_t = j | X_{t+1} = i) = p(X_{t+1} = j | X_t = i)$

Regular Markov chain A stationary Markov chain with transition probability matrix P is called regular (also called communicable) if there exists a $k > 0$ such that $[P^k]_{ij} > 0$ for all $i, j \in S$

Irreducibility A stationary Markov chain with transition probability matrix P is called irreducible if for any $i, j \in S$, $i \neq j$, there exists a $k > 0$ (possibly depending on i, j) such that $[P^k]_{ij} > 0$

Therefore irreducible MC means, it is possible to reach to any state to any other state, however it may take many steps depending on the states

Sampling methods

Question Why do they call it 'reversible'?

Homework For a reversible MC, if the initial distribution is stationary then $P(X_t = j | X_{t+1} = i) = p(X_{t+1} = j | X_t = i)$

Regular Markov chain A stationary Markov chain with transition probability matrix P is called regular (also called communicable) if there exists a $k > 0$ such that $[P^k]_{ij} > 0$ for all $i, j \in S$

Irreducibility A stationary Markov chain with transition probability matrix P is called irreducible if for any $i, j \in S$, $i \neq j$, there exists a $k > 0$ (possibly depending on i, j) such that $[P^k]_{ij} > 0$

Therefore irreducible MC means, it is possible to reach to any state to any other state, however it may take many steps depending on the states
Observation A finite Markov chain is irreducible if and only if its graph representation is a strongly connected graph.

Sampling methods

Observation For a reversible MC,

$$\sum_{j\in S} p_{ji}\pi(j) = \sum_{j\in S} p_{ij}\pi(i) = \pi(i)\sum_{j\in S} p_{ij} = \pi(i)$$

Sampling methods

Observation For a reversible MC,

$$\sum_{j\in S} p_{ji}\pi(j) = \sum_{j\in S} p_{ij}\pi(i) = \pi(i)\sum_{j\in S} p_{ij} = \pi(i)$$

Aperiodicity The period of a state $s_i \in S$ is defined as

$$gcd\{t \geq 1 : [P^t]_{ii} > 0\}.$$

Thus the period of s_i is the greatest common divisor of the set of times that the chain can return (i.e., has positive probability of returning) to s_i , given that we start with $X_0 = s_i$. If the period $s_i = 1$, then we say that the state s_i is aperiodic.

Sampling methods

Observation For a reversible MC,

$$\sum_{j \in S} p_{ji} \pi(j) = \sum_{j \in S} p_{ij} \pi(i) = \pi(i) \sum_{j \in S} p_{ij} = \pi(i)$$

Aperiodicity The period of a state $s_i \in S$ is defined as

$$gcd\{t \geq 1 : [P^t]_{ii} > 0\}.$$

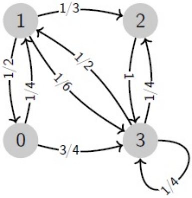
Thus the period of s_i is the greatest common divisor of the set of times that the chain can return (i.e., has positive probability of returning) to s_i , given that we start with $X_0 = s_i$. If the period $s_i = 1$, then we say that the state s_i is aperiodic.

A Markov chain is said to be aperiodic if all its states are aperiodic. Otherwise the chain is said to be periodic

Sampling methods

Homework Consider the following Markov chain:

$$P = \begin{pmatrix} 0 & \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{2} & 0 & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$



Determine the periodicity, regularity and irreduciblity property of the Markov Chain.

Sampling methods

Existence and uniqueness Any irreducible and aperiodic Markov chain has exactly one stationary distribution.

Sampling methods

Existence and uniqueness Any irreducible and aperiodic Markov chain has exactly one stationary distribution.

Example of a MC model Assume that a mosquito hops between the forehead, the left check and the right check of a person

- There are three states
- Suppose the rule is: At some time t, if the mosquito is sitting on the foreheaad then it will hop to left check at time $t+1$
- if it sits on the left check, it will stay there or move to the right check with probability 0.5 for each, and if it is on the right check then it will stay there or move to the forehead with probability 0.5 each.

Sampling methods

Existence and uniqueness Any irreducible and aperiodic Markov chain has exactly one stationary distribution.

Example of a MC model Assume that a mosquito hops between the forehead, the left check and the right check of a person

- There are three states
- Suppose the rule is: At some time t, if the mosquito is sitting on the foreheaad then it will hop to left check at time $t+1$
- if it sits on the left check, it will stay there or move to the right check with probability 0.5 for each, and if it is on the right check then it will stay there or move to the forehead with probability 0.5 each.

Then the transition matrix is

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

Sampling methods

Homework Determine the MC properties of the walk of the mosquito in the last example.

Sampling methods

Homework Determine the MC properties of the walk of the mosquito in the last example.

GOAL How should we use the MC concept for simulating a pdf?

Sampling methods

Homework Determine the MC properties of the walk of the mosquito in the last example.

GOAL How should we use the MC concept for simulating a pdf? MCMC was placed in the top 10 most important algorithms of the 20th century (SIAM news).

- $\rightarrow$ Suppose $p(\mathbf{x})$ is the given pdf
- $\rightarrow$ Define a MC on a state space whose stationary distribution is the target density

Sampling methods

A bit of history MCMC algorithm was discovered by physicists working on the atomic bomb at Los Alamos during World War II, and was first published in a chemistry journal. Metropolis et al. 1953.

Sampling methods

A bit of history MCMC algorithm was discovered by physicists working on the atomic bomb at Los Alamos during World War II, and was first published in a chemistry journal. Metropolis et al. 1953. An extension was published in the statistics literature by Hastings in 1970, but was largely unnoticed.

Sampling methods

A bit of history MCMC algorithm was discovered by physicists working on the atomic bomb at Los Alamos during World War II, and was first published in a chemistry journal. Metropolis et al. 1953. An extension was published in the statistics literature by Hastings in 1970, but was largely unnoticed. A special case (Gibbs sampling) was independently invented in 1984 in the context of Ising models and was published by Geman and Geman in 1984.

Sampling methods

A bit of history MCMC algorithm was discovered by physicists working on the atomic bomb at Los Alamos during World War II, and was first published in a chemistry journal. Metropolis et al. 1953. An extension was published in the statistics literature by Hastings in 1970, but was largely unnoticed. A special case (Gibbs sampling) was independently invented in 1984 in the context of Ising models and was published by Geman and Geman in 1984. The algorithm became well-known to the wider statistical community until 1990 (Gelfand and Smith)

Sampling methods

Math framework of MCMC algo To draw simulations from a target distribution π on a countable set S, define a stationary Markov chain or a transition matrix P which is irreducible and aperiodic such that the reversibility condition $p_{ij}\pi_i = p_{ji}\pi_j$ holds.

Sampling methods

Math framework of MCMC algo To draw simulations from a target distribution π on a countable set S, define a stationary Markov chain or a transition matrix P which is irreducible and aperiodic such that the reversibility condition $p_{ij}\pi_i = p_{ji}\pi_j$ holds.

There can be many ways (infinitely many!) such transition matrix P - these algorithms collectively known as Metropolis algorithms

Sampling methods

Math framework of MCMC algo To draw simulations from a target distribution π on a countable set S, define a stationary Markov chain or a transition matrix P which is irreducible and aperiodic such that the reversibility condition $p_{ij}\pi_i = p_{ji}\pi_j$ holds.
There can be many ways (infinitely many!) such transition matrix P - these algorithms collectively known as Metropolis algorithms

Theorem (Justification of approximating the true average by sample average) Let X_n, $n \geq 0$ be an irreducible stationary Markov chain on a discrete state space S. Suppose X_n possess a stationary distribution π. If $\phi : S \to \mathbb{R}$ is such that $\mathbb{E}_\pi[\phi(X)]$ exists then for any initial distribution μ,

$$\frac{1}{n}\sum_{k=1}^{n} \phi(X_k) \to^{\text{a.s.}} \mathbb{E}_\pi[\phi(X)]$$

Sampling methods

Metropolis Algorithms

$\rightarrow$ Suppose the chain is at some state $i \in S$ at the current time

Sampling methods

Metropolis Algorithms

- → Suppose the chain is at some state $i \in S$ at the current time
- → First step: pick a state $j \in S$ with some probability distribution for possibly moving to state j

Sampling methods

Metropolis Algorithms

- → Suppose the chain is at some state $i \in S$ at the current time
- → First step: pick a state $j \in S$ with some probability distribution for possibly moving to state j
- → The state j is called the candidate state and the distribution which is used to pick the state j is called proposal distribution

Sampling methods

Metropolis Algorithms

- → Suppose the chain is at some state $i \in S$ at the current time
- → First step: pick a state $j \in S$ with some probability distribution for possibly moving to state j
- → The state j is called the candidate state and the distribution which is used to pick the state j is called proposal distribution
- → Second step: if j is different from i then either move to the candidate state j with the designated probability or stay at the current state i

Sampling methods

Metropolis Algorithms

- → Suppose the chain is at some state $i \in S$ at the current time
- → First step: pick a state $j \in S$ with some probability distribution for possibly moving to state j
- → The state j is called the candidate state and the distribution which is used to pick the state j is called proposal distribution
- → Second step: if j is different from i then either move to the candidate state j with the designated probability or stay at the current state i

Then the entries of the transition matrix are given by

$$
\begin{aligned}
p_{ij} &= \theta_{ij}\gamma_{ij},\ i, j \in S,\ j \neq i \\
p_{ii} &= 1 - \sum_{j \neq i} p_{ij}
\end{aligned}
$$

where θ_{ij} is the probability that j is picked as a candidate state, and γ_{ij} is the probability of moving to the state j

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Homework Is there a connection to accept-reject scheme here for choosing the candidate state?

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Homework Is there a connection to accept-reject scheme here for choosing the candidate state?

Specific choices of θ_{ij}, γ_{ij} :

Independent Sampling: $\theta_{ij} = \pi(j)$ for all i, and $\gamma_{ij} \equiv 1$

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Homework Is there a connection to accept-reject scheme here for choosing the candidate state?

Specific choices of θ_{ij}, γ_{ij} :

Independent Sampling: $\theta_{ij} = \pi(j)$ for all i, and $\gamma_{ij} \equiv 1$

Metropolis–Hastings Algorithm: $\theta_{ij} = c =$ constant, and $\gamma_{ij} = \min\left\{1, \frac{\pi(j)}{\pi(i)}\right\}$

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Homework Is there a connection to accept-reject scheme here for choosing the candidate state?

Specific choices of θ_{ij}, γ_{ij} :

Independent Sampling: $\theta_{ij} = \pi(j)$ for all i, and $\gamma_{ij} \equiv 1$

Metropolis–Hastings Algorithm: $\theta_{ij} = c =$ constant, and $\gamma_{ij} = \min\left\{1, \frac{\pi(j)}{\pi(i)}\right\}$

Barker's Algorithm: $\theta_{ij} =$constant, and $\gamma_{ij} = \frac{\pi(j)}{\pi(i)+\pi(j)}$

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Homework Is there a connection to accept-reject scheme here for choosing the candidate state?

Specific choices of θ_{ij}, γ_{ij} :

Independent Sampling: $\theta_{ij} = \pi(j)$ for all i, and $\gamma_{ij} \equiv 1$

Metropolis–Hastings Algorithm: $\theta_{ij} = c =$ constant, and $\gamma_{ij} = \min\left\{1, \frac{\pi(j)}{\pi(i)}\right\}$

Barker's Algorithm: $\theta_{ij} =$constant, and $\gamma_{ij} = \frac{\pi(j)}{\pi(i)+\pi(j)}$

Independent Metropolis Algorithm: For all i, $\theta ij = p_j$, and $\gamma_{ij} = \min\left\{1, \frac{\pi(j)p(i)}{\pi(i)p_j}\right\}$

Sampling methods

Transition matrix Choose the matrix or values θ_{ij} such that the transition matrix $P = [p_{ij}]$ is irreducible

Homework Is there a connection to accept-reject scheme here for choosing the candidate state?

Specific choices of θ_{ij}, γ_{ij} :

Independent Sampling: $\theta_{ij} = \pi(j)$ for all i, and $\gamma_{ij} \equiv 1$

Metropolis–Hastings Algorithm: $\theta_{ij} = c =$ constant, and $\gamma_{ij} = \min\left\{1, \frac{\pi(j)}{\pi(i)}\right\}$

Barker's Algorithm: $\theta_{ij} =$constant, and $\gamma_{ij} = \frac{\pi(j)}{\pi(i)+\pi(j)}$

Independent Metropolis Algorithm: For all i, $\theta ij = p_j$, and $\gamma_{ij} = \min\left\{1, \frac{\pi(j)p(i)}{\pi(i)p_j}\right\}$

Homework Verifying the assumptions!