Sampling by random walk

A: Random walks on Markov chains
Sampling by random walk

Want to sample from a distribution $P$ over some finite set $\Omega$, e.g.,

- $\Omega = \{0,1\}^N$ (say, binary images on $N$ pixels)

Difficulties:

- $\Omega$ might be huge: we cannot enumerate all possible states.
- We might not be able to evaluate $P(x)$ explicitly for $x \in \Omega$ due to unknown normalization factor. But can get ratios $P(x)/P(x')$.

Solution strategy: do a random walk on $\Omega$

- Start at any state $x \in \Omega$
- Repeatedly move to a “neighboring” state, with some transition probabilities
- After a while: the distribution over the current location will be $P$

Random walks and Markov chains

Random walk on a finite state space $\Omega$

- Let $Q_t$ be the state at time $t$
  Next state $Q_{t+1}$ depends only on $Q_t$, not prior history: Markov chain

- Random walk is defined by $|\Omega| \times |\Omega|$ transition matrix

\[
M(x, x') = M_{x,x'} = \Pr(Q_{t+1} = x' | Q_t = x)
\]

- Let $\pi_t$ be the distribution of $Q_t$, so $\pi_t \in \Delta_\Omega$

\[
\pi_{t+1}(x) = \Pr(Q_{t+1} = x) = \sum_{x' \in \Omega} \Pr(Q_t = x') M_{x',x} = \sum_{x'} \pi_t(x') M_{x',x}
\]

In vector form:

\[
\pi_{t+1}^T = \pi_t^T M = \pi_{t-1}^T M^2 = \cdots = \pi_0^T M^{t+1}
\]
Stationary distribution

We say \( \pi \) is a **stationary distribution** if \( \pi^T = \pi^T M \). Such a distribution always exists.

Determine the stationary distribution of this Markov chain:

Stationary distribution: issues

(1) There may be several stationary distributions
Stationary distribution: issues

(2) The random walk may not converge to the stationary distribution, even if it is unique

Irreducible, aperiodic Markov chains

Things become easier if the Markov chain is:

- **Irreducible**: the transition graph (nodes are states, directed edges are transitions with non-zero probability) is strongly connected.

- **Aperiodic**: there exists $k > 0$ such that $M^k(x, x') > 0$ for all $x, x'$.

**Theorem.** Any irreducible, aperiodic Markov chain has a unique stationary distribution $\pi^*$. For all $x, x' \in \Omega$,

$$\lim_{t \to \infty} M^t(x, x') = \pi^*(x').$$
Figuring out the stationary distribution

**Lemma.** If $\pi$ satisfies the **detailed balance** condition:

$$\pi(x)M(x, x') = \pi(x')M(x', x) \quad \forall x, x' \in \Omega$$

then $\pi$ is a stationary distribution of $M$.

---

**B: The Gibbs sampler**
Recall: Gibbs sampler

Finite state space $\Omega = \Omega_0^N$. Want to sample from a distribution $P > 0$ on $\Omega$.

- Start with any $x \in \Omega$
- Repeat:
  - Pick a coordinate $i \in \{1, 2, \ldots, N\}$ at random
  - Resample $x_i$ from $P(X_i = x_i|\{x_j\}_{j \neq i})$

Check:

1. This is a Markov chain
2. It is irreducible and aperiodic
3. The stationary distribution is $P$
Example: Ising model

System of $N$ particles arranged in a lattice.
- Each particle has a **spin** $X_i \in \{-1, +1\}$
- Overall configuration $X = (X_1, \ldots, X_N) \in \{-1, +1\}^N$

Probability $P(x) \propto e^{-U(x)}$

**Energy** of configuration $x$:

$$U(x) = - \sum_{i, j \text{ neighbors}} J_{ij} x_i x_j - \sum_i \beta_i x_i$$

- Ferromagnetic regime: $J_{ij} > 0$
- Statistical mechanics: Local interactions $\implies$ macroscopic properties

Gibbs sampler for Ising model

Pick a particle $k \in [N]$ and resample its spin $X_k \in \{-1, +1\}$ while keeping everything else fixed.
Mixing time of a Markov chain

How many steps before the random walk gets close to the stationary distribution?

For Markov chain with transition matrix $M$, mixing time $T_{\text{mix}}(\epsilon)$ is the smallest $t$ for which

$$\max_{x \in \Omega} \| M^t(x, \cdot) - \pi^*(\cdot) \|_{TV} \leq \epsilon.$$ 

- Can show that $T_{\text{mix}}(\epsilon) \leq T_{\text{mix}}(1/4) \lceil \lg(1/\epsilon) \rceil$.

- Say $\Omega$ is of the form $\Omega_o^N$.
  - The chain is **rapidly mixing** if $T_{\text{mix}}(1/4)$ is polynomial in $N$
  - The chain is **torpidly mixing** if $T_{\text{mix}}(1/4)$ is super-polynomial (e.g. exponential) in $N$
C: Metropolis-Hastings sampler

Metropolis-Hastings walk

Want to sample from distribution $P$ on state space $\Omega$.

- We already have an irreducible, aperiodic Markov chain on it, with transition probabilities $M(x, x')$.
- But it doesn’t have the right stationary distribution. How to modify it?

- Start with any $x \in \Omega$
- Repeat:
  - Pick a new state $x' \sim M(x, \cdot)$
  - Accept it with probability

$$\min \left( \frac{P(x')M(x', x)}{P(x)M(x, x')}, 1 \right)$$

else stay at $x$
Analyzing the stationary distribution

**Theorem.** The modified chain has stationary distribution $P$. 