Markov Chains

What is a Markov Chain?

Definition: A discrete-time process *Markov Chain* is a sequence of random variables $\{X_t\}_{t\geq 0}$ with the property that, given the present state, the future and past states are independent. Formally,

$$P(X_{t+1}|X_t, X_{t-1}, \ldots, X_0) = P(X_{t+1}|X_t).$$

The Markov Chain is **time-homogeneous** if the transition probabilities do not depend on time t:

$$\forall n \in N, \quad P(X_t = y|X_{t-1} = x) = P(X_{t+m} = y|X_{t+m-1} = x)$$

i.e. transition probabilities do not depend on t. The **key idea** is MCMC is to construct a Markov Chains such that x_t converges to a desired distribution π as $t \to \infty$ and

$$\frac{1}{n}\sum_{t=1}^{n} \phi(x_t) \to \mathbb{E}_{x\sim\pi}[\phi(x)] \quad \text{as } n \to \infty.$$

What kinds of conditions are required for this to hold?

Invariant / stationary distribution

A distribution π is called **invariant** (or **stationary**) for a Markov Chain with transition kernel P if

$$\pi(y) = \int \pi(x)P(x, y)dx.$$

Intuitively, if the chain starts with distribution π, it remains in distribution π at all future times.

Time-homogeneous is not needed for invariant distribution. But it is often easier to verify in that case.

Detailed balance and reversibility

A Markov Chain with transition kernel P satisfies the **detailed balance** condition with respect to a distribution π if

$$\pi(x)P(x,y) = \pi(y)P(y,x) \quad \text{for all states } x, y.$$

If a Markov Chain satisfies detailed balance with respect to π, then π is an invariant distribution for the chain.

Detailed balance implies that the Markov Chain is **reversible** with respect to π, meaning that the process looks the same when observed forward or backward in time when started from the distribution π.

Irreducible

A Markov Chain is called **irreducible** if it is possible to get to any state from any state. Formally, for any states x and y, there exists an integer $0 \leq n < \infty$ such that

$$P^n(x, y) > 0,$$

where $P^n(x, y)$ is the n-step transition probability from state x to state y.

Aperiodicity

A Markov Chain is called **aperiodic** if it does not get trapped in cycles with fixed periods. Ensures actual convergence instead of oscillation. Formally, for any state x, the greatest common divisor of the set of integers

$$\{n \geq 1 : P^n(x, x) > 0\}$$

is 1. **Note:** If all states have a non-zero probability of remaining in the same state, the chain is aperiodic.

Convergence Conditions

For a typical MCMC algorithm to converge to target π:

- Irreducibility: Can your proposal mechanism reach all of π's support?
- Aperiodicity: Usually satisfied by having self-loops (rejection steps)
- Correct stationary distribution: Does your acceptance ratio satisfy detailed balance?

Positive recurrence

A Markov Chain is called **positive recurrent** if, starting from any state, the expected return time to that state is finite. Formally, for any state x,

$$\mathbb{E}[T_x | X_0 = x] < \infty,$$

where T_x is the return time to state x. **Note:** Positive recurrence ensures that the chain does not wander off to infinity and has a well-defined long-term behavior.

More on recurrence

- **Recurrent**: A Markov Chain is called recurrent if, starting from any state, the probability of returning to that state is 1.
- **Positive recurrence:** A Markov Chain is called positive recurrent if it is recurrent and the expected return time to any state is finite, i.e. the chain returns quickly on average.
- **Transient**:
- **Null recurrence:** A Markov Chain is called null recurrent if it is recurrent but the expected return time to any state is infinite.