18.175: Lecture 32

More Markov chains

Scott Sheffield

MIT

Outline

Outline

General setup and basic properties

Recurrence and transience

Markov chains: general definition

- Consider a measurable space $(S, \mathcal{S})$.

Markov chains: general definition

- Consider a measurable space $(S, \mathcal{S})$.
- A function $p : S \times \mathcal{S} \to \mathbb{R}$ is a **transition probability** if

Markov chains: general definition

- Consider a measurable space $(S, \mathcal{S})$.
- A function $p : S \times \mathcal{S} \to \mathbb{R}$ is a **transition probability** if
  - For each $x \in S$, $A \to p(x, A)$ is a probability measure on $S, \mathcal{S})$.

Markov chains: general definition

- Consider a measurable space $(S, \mathcal{S})$.
- A function $p : S \times \mathcal{S} \to \mathbb{R}$ is a **transition probability** if
  - For each $x \in S$, $A \to p(x, A)$ is a probability measure on $S, \mathcal{S})$.
  - For each $A \in S$, the map $x \to p(x, A)$ is a measurable function.

Markov chains: general definition

- Consider a measurable space $(S, \mathcal{S})$.
- A function $p : S \times \mathcal{S} \to \mathbb{R}$ is a **transition probability** if
  - For each $x \in S$, $A \to p(x, A)$ is a probability measure on $S, \mathcal{S})$.
  - For each $A \in S$, the map $x \to p(x, A)$ is a measurable function.
- Say that X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.

Markov chains: general definition

- Consider a measurable space $(S, \mathcal{S})$.
- A function $p : S \times \mathcal{S} \to \mathbb{R}$ is a **transition probability** if
  - For each $x \in S$, $A \to p(x, A)$ is a probability measure on $S, \mathcal{S})$.
  - For each $A \in S$, the map $x \to p(x, A)$ is a measurable function.
- Say that X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.
- How do we construct an infinite Markov chain? Choose p and initial distribution μ on $(S, \mathcal{S})$. For each $n < \infty$ write

$$P(X_j \in B_j, 0 \le j \le n) = \int_{B_0} \mu(dx_0) \int_{B_1} p(x_0, dx_1) \cdots$$

$$\int_{B_n} p(x_{n-1}, dx_n).$$

  Extend to $n = \infty$ by Kolmogorov's extension theorem.

Markov chains

- **Definition, again:** Say X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.

Markov chains

- **Definition, again:** Say X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.
- **Construction, again:** Fix initial distribution μ on $(S, \mathcal{S})$. For each $n < \infty$ write

$$P(X_j \in B_j, 0 \leq j \leq n) = \int_{B_0} \mu(dx_0) \int_{B_1} p(x_0, dx_1) \cdots \int_{B_n} p(x_{n-1}, dx_n).$$

Extend to $n = \infty$ by Kolmogorov's extension theorem.

Markov chains

- **Definition, again:** Say X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.
- **Construction, again:** Fix initial distribution μ on $(S, \mathcal{S})$. For each $n < \infty$ write

$$P(X_j \in B_j, 0 \leq j \leq n) = \int_{B_0} \mu(dx_0) \int_{B_1} p(x_0, dx_1) \cdots$$

$$\int_{B_n} p(x_{n-1}, dx_n).$$

  Extend to $n = \infty$ by Kolmogorov's extension theorem.
- **Notation:** Extension produces probability measure P_μ on sequence space $(S^{0,1,\ldots}, \mathcal{S}^{0,1,\ldots})$.

Markov chains

- **Definition, again:** Say X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.
- **Construction, again:** Fix initial distribution μ on $(S, \mathcal{S})$. For each $n < \infty$ write

$$P(X_j \in B_j, 0 \le j \le n) = \int_{B_0} \mu(dx_0) \int_{B_1} p(x_0, dx_1) \cdots$$

$$\int_{B_n} p(x_{n-1}, dx_n).$$

Extend to $n = \infty$ by Kolmogorov's extension theorem.

- **Notation:** Extension produces probability measure P_μ on sequence space $(S^{0,1,\ldots}, \mathcal{S}^{0,1,\ldots})$.
- **Theorem:** $(X_0, X_1, \ldots)$ chosen from P_μ is Markov chain.

Markov chains

- **Definition, again:** Say X_n is a **Markov chain** w.r.t. $\mathcal{F}_n$ with transition probability p if $P(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$.
- **Construction, again:** Fix initial distribution μ on $(S, \mathcal{S})$. For each $n < \infty$ write

$$P(X_j \in B_j, 0 \leq j \leq n) = \int_{B_0} \mu(dx_0) \int_{B_1} p(x_0, dx_1) \cdots$$

$$\int_{B_n} p(x_{n-1}, dx_n).$$

  Extend to $n = \infty$ by Kolmogorov's extension theorem.
- **Notation:** Extension produces probability measure P_μ on sequence space $(S^{0,1,\ldots}, \mathcal{S}^{0,1,\ldots})$.
- **Theorem:** $(X_0, X_1, \ldots)$ chosen from P_μ is Markov chain.
- **Theorem:** If X_n is any Markov chain with initial distribution μ and transition p, then finite dim. probabilities are as above.

Markov properties

- **Markov property:** Take $(\Omega_0, \mathcal{F}) = \left(S^{\{0,1,\ldots\}}, \mathcal{S}^{\{0,1,\ldots\}}\right)$, and let P_μ be Markov chain measure and θ_n the shift operator on Ω_0 (shifts sequence n units to left, discarding elements shifted off the edge). If $Y : \Omega_0 \to \mathbb{R}$ is bounded and measurable then

$$E_\mu(Y \circ \theta_n | \mathcal{F}_n) = E_{X_n} Y.$$

Markov properties

- **Markov property:** Take $(\Omega_0, \mathcal{F}) = \left(S^{\{0,1,\ldots\}}, \mathcal{S}^{\{0,1,\ldots\}}\right)$, and let P_μ be Markov chain measure and θ_n the shift operator on Ω_0 (shifts sequence n units to left, discarding elements shifted off the edge). If $Y : \Omega_0 \to \mathbb{R}$ is bounded and measurable then

$$E_\mu(Y \circ \theta_n | \mathcal{F}_n) = E_{X_n} Y.$$

- **Strong Markov property:** Can replace n with a.s. finite stopping time N and function Y can vary with time. Suppose that for each n, $Y_n : \Omega_n \to \mathbb{R}$ is measurable and $|Y_n| \leq M$ for all n. Then

$$E_\mu(Y_N \circ \theta_N | \mathcal{F}_N) = E_{X_N} Y_N,$$

where RHS means $E_x Y_n$ evaluated at $x = X_n, n = N$.

Properties

- **Property of infinite opportunities:** Suppose X_n is Markov chain and

$$P(\cup_{m=n+1}^{\infty}\{X_m \in B_m\}|X_n) \geq \delta > 0$$

on $\{X_n \in A_n\}$. Then $P(\{X_n \in A_n \, i.o.\} - \{X_n \in B_n \, i.o.\}) = 0$.

Properties

- **Property of infinite opportunities:** Suppose X_n is Markov chain and
$$P(\cup_{m=n+1}^{\infty}\{X_m \in B_m\}|X_n) \geq \delta > 0$$
on $\{X_n \in A_n\}$. Then $P(\{X_n \in A_n \, i.o.\} - \{X_n \in B_n \, i.o.\}) = 0$.
- **Reflection principle:** Symmetric random walks on $\mathbb{R}$. Have $P(\sup_{m \geq n} S_m > a) \leq 2P(S_n > a)$.

Properties

- **Property of infinite opportunities:** Suppose X_n is Markov chain and

$$P(\cup_{m=n+1}^{\infty}\{X_m \in B_m\}|X_n) \geq \delta > 0$$

on $\{X_n \in A_n\}$. Then $P(\{X_n \in A_n \, i.o.\} - \{X_n \in B_n \, i.o.\}) = 0$.

- **Reflection principle:** Symmetric random walks on $\mathbb{R}$. Have $P(\sup_{m \geq n} S_m > a) \leq 2P(S_n > a)$.
- **Proof idea:** Reflection picture.

Reversibility

- Measure μ called **reversible** if $\mu(x)p(x,y) = \mu(y)p(y,x)$ for all x, y.

Reversibility

- Measure μ called **reversible** if $\mu(x)p(x,y) = \mu(y)p(y,x)$ for all x, y.
- Reversibility implies stationarity. Implies that amount of mass moving from x to y is same as amount moving from y to x. Net flow of zero along each edge.

Reversibility

- Measure μ called **reversible** if $\mu(x)p(x,y) = \mu(y)p(y,x)$ for all x, y.
- Reversibility implies stationarity. Implies that amount of mass moving from x to y is same as amount moving from y to x. Net flow of zero along each edge.
- Markov chain called reversible if admits a reversible probability measure.

Reversibility

- Measure μ called **reversible** if $\mu(x)p(x,y) = \mu(y)p(y,x)$ for all x, y.
- Reversibility implies stationarity. Implies that amount of mass moving from x to y is same as amount moving from y to x. Net flow of zero along each edge.
- Markov chain called reversible if admits a reversible probability measure.
- Are all random walks on (undirected) graphs reversible?

Reversibility

- Measure μ called **reversible** if $\mu(x)p(x,y) = \mu(y)p(y,x)$ for all x, y.
- Reversibility implies stationarity. Implies that amount of mass moving from x to y is same as amount moving from y to x. Net flow of zero along each edge.
- Markov chain called reversible if admits a reversible probability measure.
- Are all random walks on (undirected) graphs reversible?
- What about directed graphs?

Cycle theorem

- **Kolmogorov's cycle theorem:** Suppose p is irreducible. Then exists reversible measure if and only if

Cycle theorem

- **Kolmogorov's cycle theorem:** Suppose p is irreducible. Then exists reversible measure if and only if
- Useful idea to have in mind when constructing Markov chains with given reversible distribution, as needed in Monte Carlo Markov Chains (MCMC) applications.

Cycle theorem

- **Kolmogorov's cycle theorem:** Suppose p is irreducible. Then exists reversible measure if and only if
- Useful idea to have in mind when constructing Markov chains with given reversible distribution, as needed in Monte Carlo Markov Chains (MCMC) applications.

Cycle theorem

- **Kolmogorov's cycle theorem:** Suppose p is irreducible. Then exists reversible measure if and only if
  - $p(x,y) > 0$ implies $p(y,x) > 0$
- Useful idea to have in mind when constructing Markov chains with given reversible distribution, as needed in Monte Carlo Markov Chains (MCMC) applications.

Cycle theorem

- **Kolmogorov's cycle theorem:** Suppose p is irreducible. Then exists reversible measure if and only if
  - $p(x,y) > 0$ implies $p(y,x) > 0$
  - for any loop $x_0, x_1, \ldots x_n$ with $\prod_{i=1}^n p(x_i, x_{i-1}) > 0$, we have

$$\prod_{i=1}^n \frac{p(x_{i-1}, x_i)}{p(x_i, x_{i-1})} = 1.$$

- Useful idea to have in mind when constructing Markov chains with given reversible distribution, as needed in Monte Carlo Markov Chains (MCMC) applications.

Outline

Outline

General setup and basic properties

Recurrence and transience

Query

- **Interesting question:** If A is an infinite probability transition matrix on a countable state space, what does the (infinite) matrix $I + A + A^2 + A^3 + \ldots = (I - A)^{-1}$ represent (if the sum converges)?

Query

- **Interesting question:** If A is an infinite probability transition matrix on a countable state space, what does the (infinite) matrix $I + A + A^2 + A^3 + \ldots = (I - A)^{-1}$ represent (if the sum converges)?
- **Question:** Does it describe the expected number of y hits when starting at x? Is there a similar interpretation for other power series?

Query

- **Interesting question:** If A is an infinite probability transition matrix on a countable state space, what does the (infinite) matrix $I + A + A^2 + A^3 + \ldots = (I - A)^{-1}$ represent (if the sum converges)?
- **Question:** Does it describe the expected number of y hits when starting at x? Is there a similar interpretation for other power series?
- How about e^A or $e^{\lambda A}$?

Query

- **Interesting question:** If A is an infinite probability transition matrix on a countable state space, what does the (infinite) matrix $I + A + A^2 + A^3 + \ldots = (I - A)^{-1}$ represent (if the sum converges)?
- **Question:** Does it describe the expected number of y hits when starting at x? Is there a similar interpretation for other power series?
- How about e^A or $e^{\lambda A}$?
- Related to distribution after a Poisson random number of steps?

Recurrence

- Consider probability walk from y ever returns to y.

Recurrence

- Consider probability walk from y ever returns to y.
- If it's 1, return to y infinitely often, else don't. Call y a **recurrent state** if we return to y infinitely often.