18.175: Lecture 33

Ergodic theory

Scott Sheffield

MIT

Outline

Outline

Setup

Birkhoff's ergodic theorem

Motivating problem

- Consider independent bond percolation on $\mathbb{Z}^2$ with some fixed parameter $p > 1/2$. Look at some simulations.

Motivating problem

- Consider independent bond percolation on $\mathbb{Z}^2$ with some fixed parameter $p > 1/2$. Look at some simulations.
- Let Ω be the set of maps from the edges of $\mathbb{Z}^2$ to $\{0, 1\}$, $\mathcal{F}$ the usual product σ-algebra, and $P = P_p$ the probability measure.

Motivating problem

- Consider independent bond percolation on $\mathbb{Z}^2$ with some fixed parameter $p > 1/2$. Look at some simulations.
- Let Ω be the set of maps from the edges of $\mathbb{Z}^2$ to $\{0, 1\}$, $\mathcal{F}$ the usual product σ-algebra, and $P = P_p$ the probability measure.
- Now consider an $n \times n$ box centered at 0 and ask: what fraction of the points in that box belong to an infinite clusters? Does this fraction converge to a limit (in some sense: in probability, or maybe almost surely) as $n \to \infty$?

Motivating problem

- Consider independent bond percolation on $\mathbb{Z}^2$ with some fixed parameter $p > 1/2$. Look at some simulations.
- Let Ω be the set of maps from the edges of $\mathbb{Z}^2$ to $\{0, 1\}$, $\mathcal{F}$ the usual product σ-algebra, and $P = P_p$ the probability measure.
- Now consider an $n \times n$ box centered at 0 and ask: what fraction of the points in that box belong to an infinite clusters? Does this fraction converge to a limit (in some sense: in probability, or maybe almost surely) as $n \to \infty$?
- Let $C_x = 1_{x \in \mathrm{infinitecluster}}$. If the C_x were independent or each other, then this would just be a law of large numbers question. But the C_x are not independent of each other — far from it.

Motivating problem

- Consider independent bond percolation on $\mathbb{Z}^2$ with some fixed parameter $p > 1/2$. Look at some simulations.
- Let Ω be the set of maps from the edges of $\mathbb{Z}^2$ to $\{0,1\}$, $\mathcal{F}$ the usual product σ-algebra, and $P = P_p$ the probability measure.
- Now consider an $n \times n$ box centered at 0 and ask: what fraction of the points in that box belong to an infinite clusters? Does this fraction converge to a limit (in some sense: in probability, or maybe almost surely) as $n \to \infty$?
- Let $C_x = 1_{x \in \mathrm{infinitecluster}}$. If the C_x were independent or each other, then this would just be a law of large numbers question. But the C_x are not independent of each other — far from it.
- We don't have independence. We have translation invariance instead. Is that good enough?

Motivating problem

- Consider independent bond percolation on $\mathbb{Z}^2$ with some fixed parameter $p > 1/2$. Look at some simulations.
- Let Ω be the set of maps from the edges of $\mathbb{Z}^2$ to $\{0, 1\}$, $\mathcal{F}$ the usual product σ-algebra, and $P = P_p$ the probability measure.
- Now consider an $n \times n$ box centered at 0 and ask: what fraction of the points in that box belong to an infinite clusters? Does this fraction converge to a limit (in some sense: in probability, or maybe almost surely) as $n \to \infty$?
- Let $C_x = 1_{x \in \text{infinitecluster}}$. If the C_x were independent or each other, then this would just be a law of large numbers question. But the C_x are not independent of each other — far from it.
- We don't have independence. We have translation invariance instead. Is that good enough?
- More general: C_x distributed in *some* translation invariant way, $EC_0 < \infty$. Is mean of C_x (on large box) nearly constant?

Rephrasing problem

- Let θ_x be the translation of the $\mathbb{Z}^2$ that moves 0 to x. Each θ_x induces a measure-preserving translation of Ω. Then $C_x(\omega) = C_0(\theta_{-x}(\omega))$. So summing up the C_x values is the same as summing up the $C_0(\theta_x(\omega))$ value over a range of x.

Rephrasing problem

- Let θ_x be the translation of the $\mathbb{Z}^2$ that moves 0 to x. Each θ_x induces a measure-preserving translation of Ω. Then $C_x(\omega) = C_0(\theta_{-x}(\omega))$. So summing up the C_x values is the same as summing up the $C_0(\theta_x(\omega))$ value over a range of x.
- The group of translations is generated by a one-step vertical and a one-step horizontal translation. Refer to the corresponding (commuting, P-preserving) maps on Ω as ϕ_1 and ϕ_2.

Rephrasing problem

- Let θ_x be the translation of the $\mathbb{Z}^2$ that moves 0 to x. Each θ_x induces a measure-preserving translation of Ω. Then $C_x(\omega) = C_0(\theta_{-x}(\omega))$. So summing up the C_x values is the same as summing up the $C_0(\theta_x(\omega))$ value over a range of x.
- The group of translations is generated by a one-step vertical and a one-step horizontal translation. Refer to the corresponding (commuting, P-preserving) maps on Ω as ϕ_1 and ϕ_2.
- We're interested in averaging $C_0(\phi_1^j \phi_2^k \omega)$ over a range of (j, k) pairs.

Rephrasing problem

- Let θ_x be the translation of the $\mathbb{Z}^2$ that moves 0 to x. Each θ_x induces a measure-preserving translation of Ω. Then $C_x(\omega) = C_0(\theta_{-x}(\omega))$. So summing up the C_x values is the same as summing up the $C_0(\theta_x(\omega))$ value over a range of x.
- The group of translations is generated by a one-step vertical and a one-step horizontal translation. Refer to the corresponding (commuting, P-preserving) maps on Ω as ϕ_1 and ϕ_2.
- We're interested in averaging $C_0(\phi_1^j \phi_2^k \omega)$ over a range of (j,k) pairs.
- Let's simplify matters still further and consider the one-dimensional problem. In this case, we have a random variable X and we study empirical averages of the form

$$N^{-1} \sum_{n=1}^{N} X(\phi^n \omega).$$

Examples: stationary X_j sequences

- Could take X_j i.i.d.

Examples: stationary X_j sequences

- Could take X_j i.i.d.
- Or X_n could be a Markov chain, with each individual X_j distributed according to a stationary distribution π.

Examples: stationary X_j sequences

- Could take X_j i.i.d.
- Or X_n could be a Markov chain, with each individual X_j distributed according to a stationary distribution π.
- Rotations of the circle. Say X_0 is uniform in $[0,1]$ and generally $X_j = X_0 + \alpha j$ modulo 1.

Examples: stationary X_j sequences

- Could take X_j i.i.d.
- Or X_n could be a Markov chain, with each individual X_j distributed according to a stationary distribution π.
- Rotations of the circle. Say X_0 is uniform in $[0,1]$ and generally $X_j = X_0 + \alpha j$ modulo 1.
- If $X_0, X_1, \ldots$ is stationary and $g : \mathbb{R}^{\{0,1,\ldots\}} \to \mathbb{R}$ is measurable, then $Y_k = g(X_k, X_{k+1}, \ldots)$ is stationary.

Examples: stationary X_j sequences

- Could take X_j i.i.d.
- Or X_n could be a Markov chain, with each individual X_j distributed according to a stationary distribution π.
- Rotations of the circle. Say X_0 is uniform in $[0,1]$ and generally $X_j = X_0 + \alpha j$ modulo 1.
- If $X_0, X_1, \ldots$ is stationary and $g : \mathbb{R}^{\{0,1,\ldots\}} \to \mathbb{R}$ is measurable, then $Y_k = g(X_k, X_{k+1}, \ldots)$ is stationary.
- Bernoulli shift. $X_0, X_1, \ldots$ are i.i.d. and $Y_k = \sum_{j=1}^{\infty} X_{k+j} 2^{-j}$.

Examples: stationary X_j sequences

- Could take X_j i.i.d.
- Or X_n could be a Markov chain, with each individual X_j distributed according to a stationary distribution π.
- Rotations of the circle. Say X_0 is uniform in $[0,1]$ and generally $X_j = X_0 + \alpha j$ modulo 1.
- If $X_0, X_1, \ldots$ is stationary and $g : \mathbb{R}^{\{0,1,\ldots\}} \to \mathbb{R}$ is measurable, then $Y_k = g(X_k, X_{k+1}, \ldots)$ is stationary.
- Bernoulli shift. $X_0, X_1, \ldots$ are i.i.d. and $Y_k = \sum_{j=1}^{\infty} X_{k+j} 2^{-j}$.
- Can constructed two-sided ($\mathbb{Z}$-indexed) stationary sequence from one-sided stationary sequence by Kolmogorov extension.

Examples: stationary X_j sequences

- Could take X_j i.i.d.
- Or X_n could be a Markov chain, with each individual X_j distributed according to a stationary distribution π.
- Rotations of the circle. Say X_0 is uniform in $[0,1]$ and generally $X_j = X_0 + \alpha j$ modulo 1.
- If $X_0, X_1, \ldots$ is stationary and $g : \mathbb{R}^{\{0,1,\ldots\}} \to \mathbb{R}$ is measurable, then $Y_k = g(X_k, X_{k+1}, \ldots)$ is stationary.
- Bernoulli shift. $X_0, X_1, \ldots$ are i.i.d. and $Y_k = \sum_{j=1}^{\infty} X_{k+j} 2^{-j}$.
- Can constructed two-sided ($\mathbb{Z}$-indexed) stationary sequence from one-sided stationary sequence by Kolmogorov extension.
- What if X_i are i.i.d. tosses of a p-coin, where p is itself random?

Definitions

- Say that A is **invariant** if the symmetric difference between $\phi(A)$ and A has measure zero.

Definitions

- Say that A is **invariant** if the symmetric difference between $\phi(A)$ and A has measure zero.
- Observe: class $\mathcal{I}$ of invariant events is a σ-field.

Definitions

- Say that A is **invariant** if the symmetric difference between $\phi(A)$ and A has measure zero.
- Observe: class $\mathcal{I}$ of invariant events is a σ-field.
- Measure preserving transformation is called **ergodic** if $\mathcal{I}$ is trivial, i.e., every set $A \in \mathcal{I}$ satisfies $P(A) \in \{0, 1\}$.

Definitions

- Say that A is **invariant** if the symmetric difference between $\phi(A)$ and A has measure zero.
- Observe: class $\mathcal{I}$ of invariant events is a σ-field.
- Measure preserving transformation is called **ergodic** if $\mathcal{I}$ is trivial, i.e., every set $A \in \mathcal{I}$ satisfies $P(A) \in \{0, 1\}$.
- **Example:** If $\Omega = \mathbb{R}^{\{0,1,\dots\}}$ and A is invariant, then A is necessarily in tail σ-field $\mathcal{T}$, hence has probability zero or one by Kolmogorov's $0-1$ law. So sequence is ergodic (the shift on sequence space $\mathbb{R}^{\{0,1,2,\dots\}}$ is ergodic..

Outline

Setup

Birkhoff's ergodic theorem

Outline

Setup

Birkhoff's ergodic theorem

Ergodic theorem

- Let ϕ be a measure preserving transformation of $(\Omega, \mathcal{F}, P)$. Then for any $X \in L^1$ we have

$$\frac{1}{n}\sum_{m=0}^{n-1} X(\phi^m \omega) \to E(X|\mathcal{I})$$

a.s. and in L^1.

Ergodic theorem

- Let ϕ be a measure preserving transformation of $(\Omega, \mathcal{F}, P)$. Then for any $X \in L^1$ we have

$$\frac{1}{n}\sum_{m=0}^{n-1} X(\phi^m \omega) \to E(X|\mathcal{I})$$

  a.s. and in L^1.
- Note: if sequence is ergodic, then $E(X|\mathcal{I}) = E(X)$, so the limit is just the mean.

Ergodic theorem

- Let ϕ be a measure preserving transformation of $(\Omega, \mathcal{F}, P)$. Then for any $X \in L^1$ we have

$$\frac{1}{n}\sum_{m=0}^{n-1} X(\phi^m \omega) \to E(X|\mathcal{I})$$

  a.s. and in L^1.
- Note: if sequence is ergodic, then $E(X|\mathcal{I}) = E(X)$, so the limit is just the mean.
- Proof takes a couple of pages. Shall we work through it?