Lecture 12: Subadditive Ergodic Theorem

Lecturer: *Jim Pitman* Scribe: *Songhwai Oh* <sho@eecs.berkeley.edu>

12.1 The Subadditive Ergodic Theorem

This theorem is due to Kingman and also known as Kingman's subadditive ergodic theorem [4]. Let $(\Omega, \mathcal{F}, \mathbb{P}, T)$ be a probability space with a measurable map $T : \Omega \to \Omega$ which preserves $\mathbb{P}$. The set Ω is defined such that there is a two parameter family of random variables, $(X(m,n), 0 \leq m < n < \infty)$. Each $X(m,n)$ is integrable with respect to $\mathbb{P}$ for $0 \leq m < n < \infty$. If $X(m,n)$ satisfies:

I. $X(m+1, n+1) = X(m,n) \circ T$ and

II. $X(0,n) \leq X(0,m) + X(m,n)$,

then there exists an a.s. limit

$$\lim_{n\to\infty} \frac{X(0,n)}{n} = Y \tag{12.1}$$

where $Y \in [-\infty, \infty)$ and Y is T-invariant.

Notice that Y can be $-\infty$. With additional assumptions as shown in Durrett [1], we can say more about Y. For instance, if T is ergodic, then Y is a constant.

12.2 Proof

The following proof is due to Steele [5]. The first step is to consider

$$\widetilde{X}(m,n) = X(m,n) - \sum_{k=m+1}^{n} X(k-1,k). \tag{12.2}$$

The first term is subaddtive while the second term is additive so $\widetilde{X}$ satisfies the subadditivity conditions (I) and (II). Since the subadditivity condition (II) holds,

$$\begin{aligned}
\widetilde{X}(0,2) &= X(0,2) - X(0,1) - X(1,2) \\
&\leq 0 \\
\widetilde{X}(0,3) &= X(0,3) - X(0,1) - X(1,2) - X(2,3) \\
&\leq X(0,3) - X(0,2) - X(2,3) \\
&\leq 0 \\
\ldots
\end{aligned} \tag{12.3}$$

Thus $\widetilde{X}(\cdot,\cdot) \leq 0$. Hence

$$\begin{aligned}
\widetilde{X}(0,n) &= X(0,n) - \sum_{k=1}^{n} X(k-1,k) \\
\frac{1}{n}\widetilde{X}(0,n) &= \frac{1}{n}X(0,n) - \frac{1}{n}\sum_{k=1}^{n} X(k-1,k) \\
\frac{1}{n}X(0,n) &= \frac{1}{n}\widetilde{X}(0,n) + \frac{1}{n}\sum_{k=1}^{n} X(k-1,k) \\
\frac{1}{n}X(0,n) &= \frac{1}{n}\widetilde{X}(0,n) + \frac{1}{n}\sum_{k=0}^{n-1} X(0,1)\circ T^k. \qquad (12.4)
\end{aligned}$$

By Birkhoff's ergodic theorem the last sum in (12.4) converges a.s. to a limit $\mathbb{E}(X(0,1)|\mathcal{I})$, where $\mathcal{I}$ is the T-invariant σ-field. So it is enough to show that $\widetilde{X}(0,n)/n$ has an a.s. limit and, without loss of generality, we can assume $X(\cdot,\cdot) \leq 0$.

Lemma 12.1 Let $Y = \liminf_{n\to\infty} \frac{1}{n}X(0,n)$ where $X(m,n)$ is defined as in the theorem. Then Y is a.s. invariant, i.e. $Y = Y \circ T$ a.s.

Proof: Observe that

$$\begin{aligned}
\frac{X(0,n+1)}{n+1} &= \frac{n}{n+1}\frac{X(0,n+1)}{n} \\
&\leq \frac{n}{n+1}\left\{\frac{X(0,1)}{n} + \frac{X(0,n)\circ T}{n}\right\} \quad \to \quad \frac{X(0,n)\circ T}{n} \qquad (12.5)
\end{aligned}$$

since $\frac{n}{n+1} \to 1$ and $\frac{X(0,1)}{n} \to 0$. Thus

$$\begin{aligned}
\liminf_{n\to\infty} \frac{X(0,n)}{n} &\leq \liminf_{n\to\infty} \frac{X(0,n)\circ T}{n} \\
Y &\leq Y\circ T \qquad \text{for all } \omega \in \Omega. \qquad (12.6)
\end{aligned}$$

Then for any rational α, we have $(Y > \alpha) \subset (Y\circ T > \alpha) = T^{-1}(Y > \alpha)$. But $P(Y > \alpha) = P(Y \circ T > \alpha)$ so $(Y > \alpha) = (Y \circ T > \alpha)$ a.s. for all rational α. Hence $Y = Y \circ T$ a.s. □

Now fix $\epsilon > 0$ and $M > 0$ and let $Y_M = Y \vee (-M)$ where Y is defined in Lemma 12.1 and $(x \vee y) = \max(x,y)$. Define two sets as following:

$$\begin{aligned}
B_M(L) &= \big\{\omega \in \Omega \;:\; X(0,l)(\omega) > l(Y_M(\omega)+\epsilon) \text{ for all } 1 \leq l \leq L\big\} \\
G_M(L) &= B_M(L)^C \\
&= \big\{\omega \in \Omega \;:\; \exists l \leq L \text{ s.t. } X(0,l)(\omega) \leq l(Y_M(\omega)+\epsilon)\big\}.
\end{aligned}$$

We then do a recursive cutting up of the orbit for each $\omega \in \Omega$. We decompose $[0,n)$ into three types of sets as shown in Figure 12.1. The set E is a collection of singletons σ_j where $T^{\sigma_j}(\omega) \in B_M(L)$ for $1 \leq j \leq v$. The set F is a collection of intervals $[a_i, a_i + l_i)$ for $1 \leq i \leq u$ such that $T^{a_i}(\omega) \in G_M(L)$, i.e. $X(0,l_i)\circ T^{a_i}(\omega) \leq l_i(Y_M(T^{a_i}(\omega)) + \epsilon)$. Lastly, the set H contains the remaining singletons in $[n-L, n)$. Let $w = |H|$. Notice that u, v, w are r.v.'s.

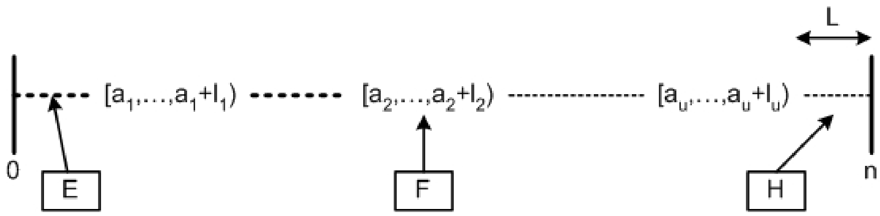

Figure 12.1: An orbit of ω, $\{\omega, T(\omega), T^2(\omega), \ldots\}$

Using the subadditivity,

$$\begin{aligned} X(0,n) &\leq \sum_{i=1}^{u} X(a_i, a_i + l_i) + \sum_{j=1}^{v} X(\sigma_j, \sigma_{j+1}) + \sum_{s=n-w}^{n} X(s, s+1) \\ &\leq \sum_{i=1}^{u} X(a_i, a_i + l_i) \\ &\leq \Big(\sum_{i=1}^{u} l_i\Big)\Big(Y_M + \epsilon\Big) \\ &\leq Y_M\Big(\sum_{i=1}^{u} l_i\Big) + n\epsilon, \end{aligned} \tag{12.7}$$

where the second inequality holds since the last two summations are less than 0 (recall that $X(\cdot,\cdot) \leq 0$).

On the other hand for all ω,

$$\sum_{i=1}^{u(\omega)} l_i(\omega) \leq n - \sum_{i=1}^{n} \mathbf{1}_{B_M(L)} \circ T^i(\omega) - L. \tag{12.8}$$

So by Birkhoff's ergodic theorem,

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{u} l_i &\leq 1 - \mathbb{E}(\mathbf{1}_{B_M(L)}|\mathcal{I}) \\ &= 1 - \mathbb{P}(B_M(L)|\mathcal{I}) \\ &= \mathbb{P}(G_M(L)|\mathcal{I}). \end{aligned} \tag{12.9}$$

Now because of the way we have defined B_M and G_M, for any fixed $M > 0$, $\mathbb{P}(G_M(L)|\mathcal{I}) \uparrow 1$ as $L \to \infty$. Combining it with (12.7) gives

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} X(0,n) &\leq Y_M \mathbb{P}(G_M(L)|\mathcal{I}) + \epsilon \\ &\xrightarrow{a.s.} Y_M + \epsilon \qquad \text{as } L \to \infty \\ &\longrightarrow Y \qquad \text{as } M \to \infty, \epsilon \to 0 \end{aligned} \tag{12.10}$$

and the proof is complete. □

12.3 Examples

12.3.1 Birkhoff's Ergodic Theorem

If $X(m,n) = \sum_{j=m}^{n-1} X \circ T^j$ where X is an integrable r.v., then $X(0,n) = X(0,m) + X(m,n)$ satisfying the conditions for the subadditive ergodic theorem. Then the conclusion from the subadditive ergodic theorem is exactly the same as that of Birkhoff's ergodic theorem.

12.3.2 Percolation

Consider an integer lattice $\mathbb{Z}^d$ in which edges $x, y \in \mathbb{Z}^d$ are connected if $|x - y| = 1$ (see Figure 12.2). Each edge e is associated with a r.v. T_e. T_e is the time taken for a message to travel across the edge e. Let $X(0,n)$ be the shortest time taken for a message to travel from $\vec{0}$ to $n\vec{v}$, where $\vec{v} = (1,0,0,\ldots,0)$. Then

$$X(0,n) = \min_{\text{path } \vec{0} \to n\vec{v}} \sum_{e \in \text{path}} T_e \tag{12.11}$$

and

$$X(m,n) = \min_{\text{path } m\vec{0} \to n\vec{v}} \sum_{e \in \text{path}} T_e. \tag{12.12}$$

Since there may be a minimum path from $\vec{0}$ to $n\vec{v}$ without visiting $m\vec{v}$,

$$X(0,n) \le X(0,m) + X(m,n). \tag{12.13}$$

Now assume T_e are independent r.v.'s with an identical distribution. So $\Omega = \{T_e, e \in E\}$, where E is a set of edges in $\mathbb{Z}^d$. Define a map $T : \Omega \to \Omega$ such that if $\omega = (t_e, e \in E)$ then $T(\omega) = (t_{e^+}, e \in E)$, where e^+ is the edge on the right of e with the same orientation (see Figure 12.2). The subadditive ergodic theorem tells that if T is ergodic then there is a constant limit $Y = \lim_{n\to\infty} \frac{X(0,n)}{n}$, meaning that a message moves at a constant speed. For more about percolation, see Grimmett [2].

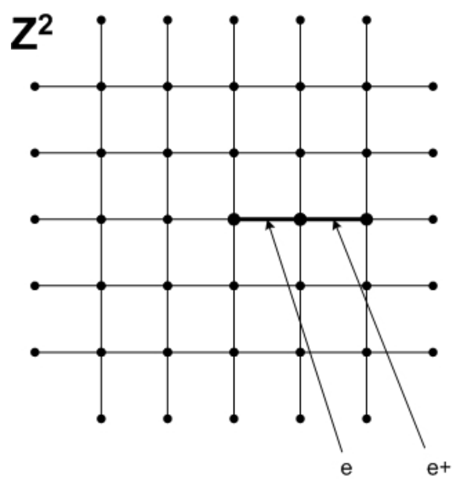

Figure 12.2: Z^2 lattice

12.3.3 Longest Increasing Subsequence

The problem of studying the behavior of the longest increasing subsequence of a random permutation is known as a problem of Ulam and Hammersley[3] was first to study this problem analytically using the subadditive ergodic theorem.

Let σ_n be a permutation of $\{1, 2, \ldots, n\}$ and assume that all $n!$ permutations are equally likely. An increasing subsequence of a permutation σ_n is a subsequence $(\sigma(k_1), \sigma(k_2), \ldots, \sigma(k_l))$ such that $1 \le k_1 < k_2 < \cdots < k_l \le n$ and $\sigma_n(k_1) < \sigma_n(k_2) < \cdots < \sigma_n(k_l)$, where l is the length of the increasing subsequence. Let L_n be the length of the longest increasing subsequence. For example, if we are given a set $\{1, 2, 3, 4, 5\}$ and its permutation $\sigma_n(1) = 1, \sigma_n(2) = 3, \sigma_n(3) = 5, \sigma_n(4) = 2, \sigma_n(5) = 4$, then increasing subsequences are $(1, 2, 4), (1, 3, 5), (1, 3, 4), (1, 2), \ldots$. Thus $L_5 = 3$.

The main difficulties with this problem is that it is hard to give any reasonable expression for either $\mathbb{E}L_n$ or $Var(L_n)$ and it is nontrivial to find the asymptotic behavior of L_n as n tends infinity. Hammersley [3] has shown the existence of an a.s. limit of $L_n/\sqrt{n}$ by ingeniously transforming the problem into a unit rate Poisson process on a two dimensional plane and using the subadditive ergodic theorem. Later it has been shown that $\lim_{n\to\infty} L_n/\sqrt{n} = 2$ a.s. See Durrett [1] for the references on this topic.

Instead of working with random permutations, we may as well work with a sequence of i.i.d. r.v.'s with a continuous distribution. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.'s with uniform distribution on $[0, 1]$. Let $\hat{L}_n$ be the length of the longest increasing subsequence in $X_1, X_2, \ldots, X_n$. Then $\hat{L}_n =_d L_n$ since there is no tie a.s. (each X_n has a continuous distribution) and the ranks of $X_1, X_2, \ldots, X_n$ corresponds to a permulation of $\{1, 2, \ldots, n\}$. By using i.i.d. r.v.'s we have randomized the vertical scale as shown in Figure 12.3.

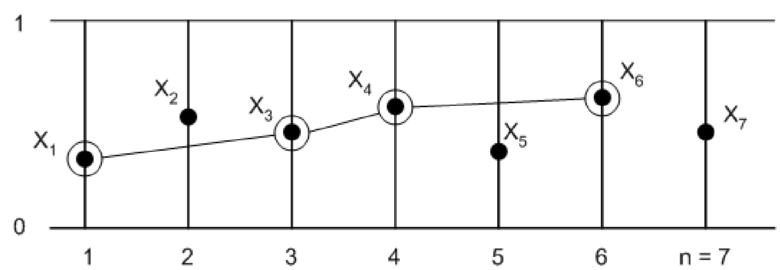

Figure 12.3: An instance of $X_1, X_2, \ldots, X_7 \sim Unif[0, 1]$ i.i.d. and its longest increasing subsequence (in circle), $L_7 = 4$

Now consider a Poisson point process in $(0, \infty) \times (0, \infty)$ with unit intensity per unit area. See Figure 12.4. Take an $k \times k$ box. Let M_k^* be the length of the longest inceasing subsequence of points in the $k \times k$ box. Let N_k be the number of points in the $k \times k$ box. Say $N_k = n$. The vertical values are i.i.d. uniform on $[0, k]$ like the previous example with uniform i.i.d r.v.'s. Hence $(M_k^* | N_k = n) = L_n$ in distribution. By superadditivity, $M_{k+m}^* \ge M_k^* + M_m^* \circ T^k$, where T^k shifts the origin from $(0, 0)$ to (k, k) and preserves the measure (see Figure 12.4). Now apply the subadditive ergodic theorem to $-M_n^*$ and conclude that

$$\frac{M_k^*}{k} \xrightarrow{a.s.} Y \in [0, \infty], \tag{12.14}$$

where Y is invariant relative to T^k and since T is ergodic Y is constant.

Lastly, let's understand the relationship between L_n and $\sqrt{n}$. We have seen that given $N_k = n$, M_k^* behaves like L_n. Now notice that for large k, $N_k = n$ is approximately k^2, the average number of points in $k \times k$ box, so $n \approx k^2$. Hence $M_k^*/k \approx L_n/\sqrt{n}$.

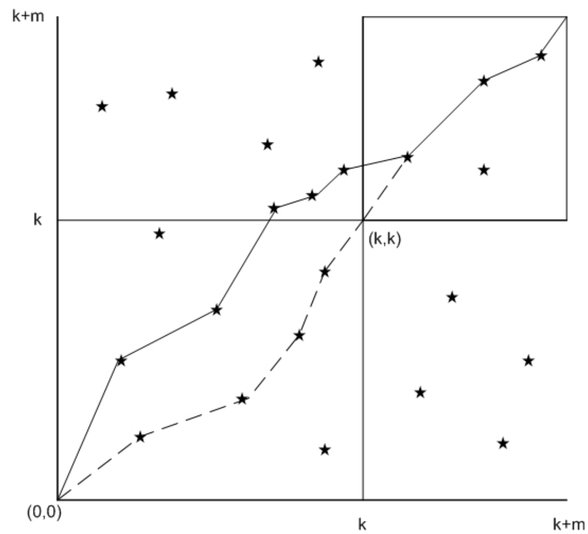

Figure 12.4: A longest increasing subsequence from a Poisson process. The solid line shows the longest increasing subsequence in M^*_{k+m} while the dotted line shows the longest increasing subsequence in M^*_k and $M^*_m \circ T^k$.

References

[1] R. Durrett. *Probability: Theory and Examples*. Duxbury Press: Belmont, CA, 1996.

[2] G. Grimmett. *Percolation*. Springer-Verlag, 1989.

[3] J.M. Hammersley. A few seedlings of research. *Proc. 6 th Berkeley Symp. Math. Stat. Prob., U. of California Press*, pages 345–394, 1972.

[4] J.F.C. Kingman. Subadditive ergodic theory. *Ann. Probab.*, pages 883–909, 1976.

[5] J.M. Steele. Kingman's subadditive ergodic theorem. *Annales de l'Institute Henri Poincare*, pages 93–98, 1989.