The Arzelà-Ascoli Theorem

John Duchi: notes for Statistics 300b

February 20, 2017

1 Introduction

In asymptotic statistics, uniform central limit theorems provide some of the most powerful generalizations of the general limit theorem to infinite dimensional problems. Any convergence results for infinite dimensional processes necessarily rely on some type of limiting compactness argument, so that random vectors and processes are well-defined. To that end, it is important to understand and quantify compactness of collections of functions.

The Arzelà-Ascoli theorem is a foundational result in analysis, and it gives necessary and sufficient conditions for a collection of continuous functions to be compact. Recall that a metric space (T, d) is compact if every open cover of T has a finite subcover, and that this is equivalent to sequential compactness, that is, that any sequence $\{t_n\} \subset T$ has a convergent subsequence with limit in T. We also recall that a set is *relatively compact* if its closure is compact. So, for example, if every sequence $\{t_n\} \subset T$ for a metric space (T, d) has a convergent subsequence $t_{n(k)}$ with limit $t_\infty = \lim_{k\to\infty} t_{n(k)}$, but for which t_∞ does not necessarily lie in T, then T is relatively compact (as $\operatorname{cl} T$ will be compact).

2 Equicontinuity

Let (T, d) be a compact metric space, where d is a metric on T, and let $\mathcal{C}(T, \mathbb{R})$ denote the collection of continuous functions $f : T \to \mathbb{R}$, where the metric is the usual supremum norm, that is,

$$\|f - g\|_\infty = \sup_{t\in T} |f(t) - g(t)| \ \text{ for } f, g \in \mathcal{C}(T, \mathbb{R}).$$

Let $\mathcal{F} \subset \mathcal{C}(T, \mathbb{R})$. Then the *modulus of continuity* of f is

$$\omega_f(\delta) := \sup_{s,t\in T} \{|f(t) - f(s)| : d(s, t) < \delta\}.$$

A function f is uniformly continuous if and only if $\lim_{\delta\to0} \omega_f(\delta) = 0$. Making this uniform over a collection of functions yields the following definition.

Definition 2.1. *A collection of functions $\mathcal{F}$ is* uniformly equicontinuous on T *if*

$$\lim_{\delta\to0} \sup_{f\in\mathcal{F}} \omega_f(\delta) = 0.$$

Equivalently, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\sup_{d(s,t)<\delta} |f(s) - f(t)| \le \epsilon \ \text{ for all } f \in \mathcal{F}.$$

Noting that the supremum above extends over all pairs $s, t \in T$ that are close together, we see that the collection of functions $\mathcal{F}$ is uniformly continuous with the *same* "degree" of uniform continuity for each $f \in \mathcal{F}$. Of course, any continuous function on a compact set T is uniformly continuous, so Definition 2.1 extends this in a uniform way to collections of functions.

One simple consequence of Definition 2.1 is that if a function class $\mathcal{F}$ is bounded at some $t_0 \in T$, then it is bounded for all $t \in T$. Indeed, suppose that $M \geq \sup_{f \in \mathcal{F}} |f(t_0)|$. Fix any $\epsilon > 0$ and choose $\delta > 0$ such that $\sup_{d(s,t)<\delta} |f(s) - f(t)| < \epsilon$. The collection of open sets $U_t := \{s \in T : d(s,t) < \delta\}$ covers T, and hence has a finite subcover $\{U_{t_i}\}_{i=1}^N$. Then for any $f \in \mathcal{F}$ and $t \in T$, there is some chain $t \to t_{i_1} \to t_{i_2} \to \ldots \to t_{i_k} \to t_0$ with $k \leq N$ and $d(t, t') < \delta$ for each pair on the chain. We thus have

$$\begin{aligned} |f(t)| \leq |f(t) - f(t_0)| + |f(t_0)| &\leq |f(t) - f(t_1)| + |f(t_k) - f(t_0)| + |f(t_0)| + \sum_{j=1}^{k-1} |f(t_j) - f(t_{j+1})| \\ &\leq \epsilon + \epsilon + |f(t_0)| + k\epsilon \leq |f(t_0)| + (N+2)\epsilon. \qquad (1) \end{aligned}$$

As f is arbitary, this shows the pointwise boundedness.

3 The theorem

It turns out that Definition 2.1 is precisely what is needed to show that collections of continuous functions are compact.

Theorem 1. *Let (T, d) be a compact metric space. Then $\mathcal{F} \subset \mathcal{C}(T, \mathbb{R})$ is relatively compact if and only if it is uniformly equicontinuous and for some $t_0 \in T$ we have $\sup_{f \in \mathcal{F}} |f(t_0)| < \infty$.*

An equivalent statement to the theorem, which is what we in fact prove, is that the two statements

(i) Any sequence $\{f_k\} \subset \mathcal{F} \subset \mathcal{C}(T, d)$ has a convergent subsequence in the supremum norm

(ii) The collection $\mathcal{F}$ is uniformly equicontinous and pointwise bounded, i.e. $\sup_{f \in \mathcal{F}} |f(t)| < \infty$ for all $t \in T$

are equivalent. (In statement (i), that $\sup_{f \in \mathcal{F}} |f(t_0)| < \infty$ is equivalent to $\sup_{f \in \mathcal{F}} |f(t)| < \infty$ for all $t \in T$ is a consequence of our calculation (1).)

Proof We begin by proving that (ii) implies (i). Let the collection $\mathcal{F}$ be uniformly equicontinuous and pointwise bounded. Our proof follows a diagonalization argument. Let $\{f_k\}_{k=1}^\infty \subset \mathcal{F}$ be a sequence of functions. As T is compact it is separable (take finite covers of radius 2^{-n} for $n \in \mathbb{N}$, pick a point from each open set in the cover, and let $n \to \infty$). Let T^0 denote a countable dense subset of T and fix an enumeration $\{t_1, t_2, \ldots\}$ of T^0. For each i define

$$F_i := \{f_k(t_i)\}_{k=1}^\infty,$$

each of which is a bounded subset of $\mathbb{R}$ by the pointwise boundedness assumption.

We may then define a nested subsequence of functions $\{f_k\} \supset \{f_k^1\} \supset \{f_k^2\} \supset \ldots$ as follows. Let $\{f_k^1\} \subset \{f_k\}$ be any subsequence of f_k for which $f_k(t_1)$ converges as $k \to \infty$, which exists because F_1 is bounded. Similarly, $\{f_k^2\} \subset \{f_k^1\}$ be any subsequence of $\{f_k^1\}$ for which $f_k^1(t_2)$ converges; generally, we let $\{f_k^n\} \subset \{f_k^{n-1}\}$ be a further subsequence of $\{f_k^{n-1}\}$ such that $f_k^{n-1}(t_n)$ converges. Now, define the diagonal subsequence $f_{n(k)} = f_k^k$. We show that $f_{n(k)}$ is Cauchy for the $\|\cdot\|_\infty$-norm on T. Fix $\epsilon > 0$, and let $\delta > 0$ be such that $\sup_{f \in \mathcal{F}} \omega_f(\delta) < \epsilon$. The collection of open sets

$U_i := \{t \in T : d(s, t_i) < \delta\}$ for $t_i \in T^0$ covers T, so that there is a finite subcover generated (w.l.o.g. by re-ordering the t_i) by $t_1, \ldots, t_N \in T^0$. Let $t \in T$ be arbitrary and note that $d(t, t_i) < \delta$ for some t_i with $i \leq N$. Then for any k, k' we have

$$\begin{aligned} |f_{n(k)}(t) - f_{n(k')}(t)| &\leq |f_{n(k)}(t) - f_{n(k)}(t_i)| + |f_{n(k)}(t_i) - f_{n(k')}(t_i)| + |f_{n(k')}(t_i) - f_{n(k')}(t)| \\ &\leq \epsilon + |f_{n(k)}(t_i) - f_{n(k')}(t_i)| + \epsilon \end{aligned}$$

by the uniform equicontinuity. But of course, there only finitely many t_i, and so for large enough k, k', we have

$$\max_{i \in \{1, \ldots, N\}} \left| f_{n(k)}(t_i) - f_{n(k')}(t_i) \right| \leq \epsilon.$$

These bounds may be taken independent of t, so that $\lim_{k,k' \to \infty} \|f_{n(k)} - f_{n(k')}\|_\infty = 0$. The sequence $\{f_{n(k)}\}_{k \in \mathbb{N}} \subset \mathcal{C}(T, \mathbb{R})$ is thus Cauchy for the supremum norm on T and must converge to some continuous function.

We now show how the existence of convergent subsequences (i) implies equicontinuity and pointwise boundedness (ii). Note that the completion of $\mathcal{F}$ in the sup-norm is compact by the equivalence of sequential compactness and compactness. First, the pointwise boundedness of $\mathcal{F}$ is immediate: if $\mathcal{F}$ is compact, then $F_t := \{f(t)\}_{f \in \mathcal{F}}$ must be bounded by the corresponding compactness results on $\mathbb{R}$ (certainly all sequences within F_t must have a convergent subsequence). We would like to show the unifrom equicontinuity result that $\lim_{\delta \to 0} \sup_{f \in \mathcal{F}} \omega_f(\delta) = 0$. First, note that $f \mapsto \omega_f(\delta)$ is Lipschitz in f: we have by the triangle inequality that

$$\omega_f(\delta) - \omega_g(\delta) = \sup_{d(s,t) < \delta} |f(t) - f(s)| - \sup_{d(s,t) < \delta} |g(t) - g(s)| \leq \sup_{d(s,t) < \delta} |f(t) - g(t) + f(s) - g(s)| \leq 2 \|f - g\|_\infty,$$

so that $f \mapsto \omega_f(\delta)$ is 2-Lipschitz in $\|\cdot\|_\infty$-norm for any fixed $\delta > 0$. Moreover, as T is compact, we must have $f \in \mathcal{C}(T, \mathbb{R})$ uniformly continuous, so that $\omega_f(n^{-1}) \to 0$ as $n \to \infty \to 0$. That is, the pointwise limit of the sequence of functions $h_n(f) := \omega_f(n^{-1})$ as $n \to \infty$ is the zero function. But the collection $\{h_n\}$ is (uniformly) Lipschitz, and the standard result that the pointwise limit of uniformly Lipschitz functions on a compact set is the same as the uniform limit gives that $\sup_{f \in \mathcal{F}} h_n(f) = \sup_{f \in \mathcal{F}} \omega_f(n^{-1}) \to 0$. □