THE ARZELÀ-ASCOLI THEOREM AND ITS APPLICATIONS

ADAM QUINN JAFFE

We provide a statement and proof of the Arzelà-Ascoli theorem and then explore some applications to ordinary differential equations and functional analysis. These notes are based on the material of MATH 171 at Stanford during Spring 2017, taught by Steven Kerckhoff.

Compactness is among the most useful properties that a metric space can possess, but it is often to hard to check for it directly. For this reason, there are many equivalent characterizations of compactness that may be easier to check in various contexts: sequential compactness, complete and totally bounded, etc. The Arzelà-Ascoli Theorem provides another useful characterization in functional analysis: the closure of a set of continuous functions is compact if and only if it is equicontinuous and uniformly bounded.

1. Introduction

Let us first define our basic object of study:

Definition 1.1. Let $C(I) = \{f : I \to \mathbb{R} \mid f \text{ continuous}\}$ denote the set of continuous functions from the closed interval $I = [0, 1]$ into the real numbers $\mathbb{R}$. Let the *sup distance* $d : C(I) \times C(I) \to \mathbb{R}$ be the function given by $d(f, g) = \sup_{x \in I} |f(x) - g(x)|$.

For brevity's sake, we simply state the following without proof:

Theorem 1.2. The space $(C(I), d)$ is a complete metric space.

We are now equipped to tackle the main theorem and proof. Let us precisely state the conditions we will later use:

Definition 1.3. A set Φ is called *uniformly bounded* if there exists a bound $B > 0$ such that $|f(x)| \leq B$ holds for all $x \in I$ and $f \in \Phi$.

Definition 1.4. A set Φ is called *equicontinuous* if, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$ for all $x, y \in I$ and $f \in \Phi$.

The main theorem can then be stated as follows:

Theorem 1.5 (Arzelà-Ascoli)**.** If $\Phi \subseteq C(I)$ is equicontinuous and uniformly bounded, then the closure $\overline{\Phi}$ is compact.

2. Necessity of Hypotheses

Before proving the main theorem, we illustrate the importance of both of the hypotheses listed above. To see why the uniform boundedness condition is necessary, consider the following example:

Example 2.1. Let $\Phi = \{f_n(x) = n \mid n \in \mathbb{N}\}$. It is clear that Φ is not uniformly bounded, since we can always pick the integer $N = \lceil B \rceil$ so that $f_N(0) \geq B$. However, we see that Φ is equicontinuous, since $|f_n(x) - f_n(y)| = 0$ for all $x, y \in I$

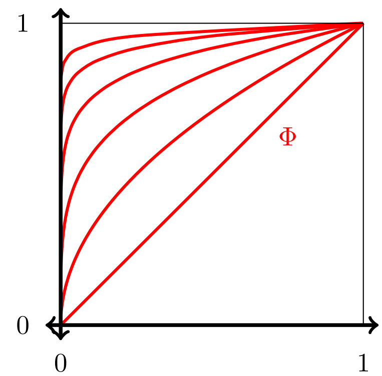

FIGURE 1. A visualization of Example 2.2

and $n \in \mathbb{N}$. That means, for any $\varepsilon > 0$, we can simply take $\delta = 1$ and the condition is met. Finally, we will show that the conclusion fails in this example, i.e. that $\bar{\Phi}$ is not compact: Consider the sequence $\{f_n(x)\}_{n=1}^{\infty}$ and observe that $d(f_m, f_n) = |m - n| \geq 1$ for any distinct indices. Thus, this sequence cannot have a convergent subsequence, for it is not even Cauchy: There is no N that has $d(f_m, f_n) < 1/2$ for all $m, n \geq N$.

To see why condition equicontinuity is necessary, we consider another example:

Example 2.2. Let $\Phi = \{f_n(x) = x^{1/n} \mid n \in \mathbb{N}\}$. In this case, uniform boundedness is satisfied, since $0 \leq f_n(x) \leq 1$ for all $x \in I$ and $n \in \mathbb{N}$, i.e. we can take $B = 1$ as the uniform bound. Equicontinuity, however, is not met by this example: For any $\varepsilon > 0$, assume that there was a $\delta > 0$ satisfying the property. Then using the fact that the $\{\varepsilon^n\} \to 0$ as $n \to \infty$, we can pick an N such that $\varepsilon^N < \delta$. Then we use the assumption on the function f_N to conclude that $|\varepsilon^N - 0| < \delta$ implies $\varepsilon = |(\varepsilon^N)^{1/N} - 0^{1/N}| < \varepsilon$. This is a contradiction, hence Φ is not equicontinuous. We can now verify that the conclusion indeed fails: For $x = 0$, we have $\lim_{n\to\infty} f_n(x) = 0$; for $x > 0$, we have $\lim_{n\to\infty} f_n(x) = 1$. Therefore, $\{f_n\}$ converges pointwise to the function f, defined as:

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } 0 < x \leq 1 \end{cases}$$

This can be visualized with the help of Figure 1, which depicts $f_1, f_2, f_4, \ldots f_{32}$. Since uniform convergence implies pointwise convergence, we have shown that the sequence $\{f_n\}$ (and hence any subequence of it) must converge to this same limit. However, f is not continuous, so Φ has no convergent subsequences.

Now we can proceed to the proof of the theorem.

3. PROOF OF THEOREM 1.5

The main idea of the proof is quite straightforward. We will create a finite lattice with a high enough "resolution" so that piecewise linear functions passing through the points on this lattice make a sufficiently accurate approximation of the functions in Φ. Since there are only finitely many functions that can be defined on this lattice, we will have shown that Φ can indeed be covered by finitely many balls of an arbitrarily small fixed radius.

Suppose that $\Phi \subseteq C(I)$ is a subset satisfying both uniform boundedness and equicontinuity. Since, $C(I)$ is complete, and since any closed subset of a complete space must itself be complete, we conclude that $\bar{\Phi}$ is complete. Thus, since a set is compact if and only if it is complete and totally bounded, it only remains to prove that $\bar{\Phi}$ is totally bounded. To do so, we will show that Φ is totally bounded and invoke the following proposition:

Claim 3.1. If $X \subseteq M$ is totally bounded, then $\bar{X}$ is totally bounded.

Proof. Let $\varepsilon > 0$ be arbitrary. Now use the total boundedness of X to create points $x_1, \dots x_k$ satisfying $X = B_{\varepsilon/2}(x_1) \cup \dots \cup B_{\varepsilon/2}(x_k)$. We now show that $\bar{X} = B_\varepsilon(x_1) \cup \dots \cup B_\varepsilon(x_k)$. Let a be a limit point of X so that the sequence $\{a_n\}$ in X goes to a. Pick some N so that $d(a_N, a) < \varepsilon/2$ and let x_j be the center of a ball containing a_N. Then use the triangle inequality to conclude:

$$d(a, x_j) \le d(a, a_n) + d(a_n, x_j) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Therefore, a is contained in the open ball $B_\varepsilon(x_j)$. This shows that any limit point of X is contained in one of these open balls, so $\bar{X} = B_\varepsilon(x_1) \cup \dots \cup B_\varepsilon(x_k)$. Hence, $\bar{X}$ is totally bounded. □

Now we must prove that Φ is totally bounded. To do so, let $\varepsilon > 0$ be arbitrary. Now take a uniform bound B, and divide the interval $[-B, B]$ into m equally-sized subintervals of length strictly less than ε. This gives rise to the sequence $-B = y_1, y_2, \dots y_{m-1}, y_m = B$ such that $y_{j+1} - y_j < \varepsilon$ for all $j = 1, \dots m-1$. Likewise, take δ as in equicontinuity, and divide the interval $[0, 1]$ into n equally-sized subintervals of length less than δ. This gives us the sequence $0 = x_1, x_2, \dots x_{n-1}, x_n = 1$ such that $x_{i+1} - x_i < \delta$ for all $i = 1, \dots n-1$.

Note now that we have constructed a lattice $\{(x_i, y_j) \mid i = 1, \dots n,\ j = 1, \dots m\}$ which has exactly mn points. Moreover, there are exactly m^n possible piecewise linear functions which have the property that, for any i, there exists some j such that $\varphi(x_i) = y_j$. We now show that Φ is covered by the set of balls centered at each of these functions.

For any $f \in \Phi$, we assign an approximation φ in the following way: For each i, define the value of $\varphi(x_i)$ to be y_j, where y_j is the quantized output minimizing the error distance $|f(x_i) - y_j|$ (breaking ties by, say, always choosing the smaller value); then connect the values of $\varphi(x_i)$ together by a linear function on each interval; See Figure 2 for an example of this procedure. It is now clear that φ is a piecewise linear function, and that

$$|\varphi(x_k) - f(x_k)| \le \frac{\varepsilon}{2}$$

holds for all indices k.

We now show that f is contained in the open ball of radius $\frac{7}{2}\varepsilon$ centered at φ, which amounts to computing the distance $d(f, \varphi) = \sup_{x \in I} |f(x) - \varphi(x)|$. First note that we have the inequalities $|\varphi(x_k) - f(x_k)| \le \frac{\varepsilon}{2}$ and $|\varphi(x_{k+1}) - f(x_{k+1})| \le \frac{\varepsilon}{2}$ immediately by the construction of φ, and that we have $|f(x_k) - f(x_{k+1})| < \varepsilon$ by equicontinuity. Adding these together and using the triangle inequality gives the bound: $|\varphi(x_k) - \varphi(x_{k+1})| < 2\varepsilon$ for any x_k.

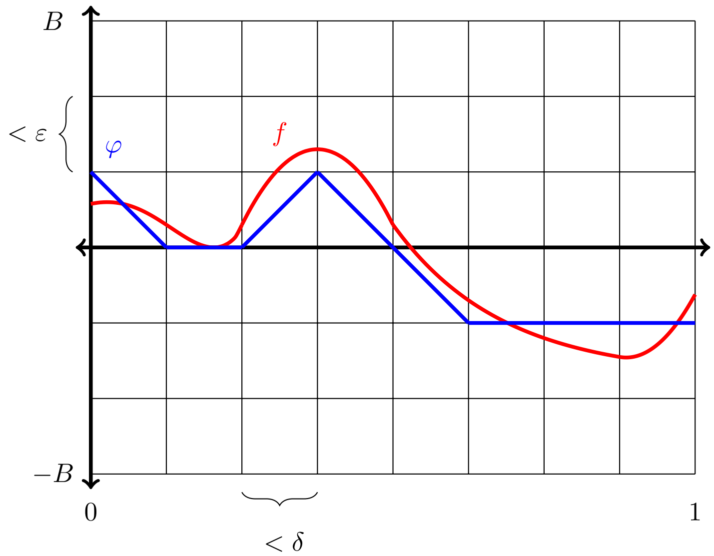

FIGURE 2. An example of the discretization procedure used in the proof of the main result.

Now, for any given x, we can choose an index k such that $x_k \leq x \leq x_{k+1}$ simply by taking the largest k with $x_k \leq x$. Then, φ is linear on the interval $[x_k, x_{k+1}]$ so it has a constant slope c. This allows us to derive the bound:

$$\begin{aligned}
|\varphi(x_k) - \varphi(x)| &= |\varphi(x_k - x)| \\
&= |c| \cdot |x_k - x| \\
&\leq |c| \cdot |x_k - x_{k+1}| \\
&= |\varphi(x_k - x_{k+1})| \\
&= |\varphi(x_k) - \varphi(x_{k+1})|
\end{aligned}$$

Therefore, we have established $|\varphi(x_k) - \varphi(x)| \leq |\varphi(x_k) - \varphi(x_{k+1})| < 2\varepsilon$. Lastly, recall that $|f(x) - f(x_k)| < \varepsilon$ by equicontinuity and that $|f(x_k) - \varphi(x_k)| < \frac{\varepsilon}{2}$ by the construction of φ. Combine this with the result just proved that $|\varphi(x_k) - \varphi(x)| < 2\varepsilon$ to conclude $|f(x) - \varphi(x)| < \frac{7}{2}\varepsilon$. Hence, $d(f, \varphi) = \sup_{x \in I} |f(x) - \varphi(x)| \leq \frac{7}{2}\varepsilon$, so we have shown that we can cover Φ by a finite number of balls of arbitrarily small size, hence Φ is totally bounded. This proves the desired result, and completes the proof that $\bar{\Phi}$ is compact.

4. Applications

As stated above, the theorem has many applications across analysis. We now describe a few such results, all of which concern families of functions which are, in some intuitive sense, "close together". These examples draw on much terminology and background from other fields, so they are meant to be expository rather than formal—just to convey the power of the Arzelà-Ascoli Theorem in these varied contexts.

Theorem 4.1. Let K be any continuous function from $I \times I$ to $\mathbb{R}$, and define the linear operator $T_K : C(I) \to C(I)$ as:

$$(T_K f)(x) = \int_0^1 K(x,y)f(y)dy$$

Then T_K is a compact operator.

Theorem 4.2. Let f be a function from $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$ which is continuous at $(0, x_0)$. Then the differential equation $\frac{dx}{dt} = f(t,x)$ with initial value $x(0) = x_0$ has a local solution, i.e. one defined on $[0, \varepsilon]$ for some $\varepsilon > 0$.

References

[1] Richard Johnsonbaugh, W.E. Pfaffenberger. *Foundations of Mathematical Analysis*. Dover Publications, Inc., 2002.