Topological preliminaries

The space $C(\mathcal{K})$ of continuous functions on $\mathcal{K}$ is a separable Banach space (complete, normed, linear space) when endowed with the supremum norm $\|f\| = \sup_{x\in\mathcal{K}} |f(x)|$.

Remark 1. Recall the following. A topological space is compact if any cover of it by open sets admit a finite sub-cover. A topological space is separable if it has a countable dense set. A compact metric space is separable (Proof. For every $n \geqslant 1$, by compactness, there exists a finite cover of balls of radius $1/n$. The set of all the centers of such balls is countable and dense.)

The separability of $C(\mathcal{K})$ is a consequence of the Stone-Weierstrass theorem. Recall that the Weierstrass theorem says that polynomials are dense among continuous functions on $[0, 1]$ (there exists a nice probabilistic proof of this fact).

Theorem 2. (Stone-Weierstraß) *Let $\mathcal{A} \subseteq C(\mathcal{K})$ be an algebra which separates the points of $\mathcal{K}$ (i.e. such that for any $x, y \in \mathcal{K}$ there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$) and such that $1 \in \mathcal{A}$. Then $\mathcal{A}$ is dense in $C(\mathcal{K})$.*

Proof. We need to prove that $\bar{\mathcal{A}} = C(\mathcal{K})$ or equivalently that we can approximate arbitrary function $f \in C(\mathcal{K})$ by an algebraic construction involving only functions in $\mathcal{A}$. We will prove that for any f and any $\varepsilon > 0$ we can find $g \in \bar{\mathcal{A}}$ such that $f < g < f + \varepsilon$ this will be enough to conclude since then $\|f - g\| \leqslant \varepsilon$. The construction has three parts.

First we need to show that if $f, g \in \mathcal{A}$ then $\max(f, g) \in \bar{\mathcal{A}}$. Since $\max(f, g) = (|f + g| + |f - g|)/2$ it is enough to prove that $f \in \mathcal{A} \Rightarrow |f| \in \bar{\mathcal{A}}$. But by Weierstrass approximation theorem, for every $\varepsilon > 0$ there exists a polynomial P_ε: $[-1, 1] \to \mathbb{R}$ such that $\sup_{-1\leqslant x\leqslant 1} ||x| - P_\varepsilon(x)| \leqslant \varepsilon$ which give us that $\|f - \|f\| P_\varepsilon(f/\|f\|)\| \leqslant \varepsilon$ and since ε is arbitrary that $|f| \in \bar{\mathcal{A}}$.

Second we use the separation property of $\mathcal{A}$. Fix $x \in \mathcal{K}$. For any $y \in \mathcal{K}$, $y \neq x$ there exists a function $h_{x,y} \in \mathcal{A}$ such that $h_{x,y}(x) \neq h_{x,y}(y)$ and we can always choose this function so that $h_{x,y}(x) = f(x) + \varepsilon/2$ and $h_{x,y}(y) = f(y) - \varepsilon/2$ (by a linear transformation using the fact that $1 \in \mathcal{A}$). Let $U_{x,y} = \{z \in \mathcal{K}: h_{x,y}(z) < f(z) + \varepsilon\}$ then $\{x, y\} \in U_{x,y}$ and $\cup_{y\neq x} U_{x,y} = \mathcal{K}$ so $\{U_{x,y}\}_{y\neq x}$ is a cover of $\mathcal{K}$ by open sets (since $h_{x,y}$ and f are continuous). By compactness we can extract a finite sub-cover $\{U_{x,y_i}\}_{i=1,\ldots,n}$ and for any $z \in \mathcal{K}$ we have $h_{x,y_i}(z) < f(z) + \varepsilon$ for any $i = 1, \ldots, n$. Then $g_x(z) = \max_i (h_{x,y_i}(z)) \in \bar{\mathcal{A}}$ and $g_x < f + \varepsilon$ on $\mathcal{K}$.

Third step is to consider the open sets $V_x = \{z \in \mathcal{K}: g_x(z) > f(z)\}$. By construction $x \in V_x$ since $g_x(x) = f(x) + \varepsilon/2$. So $\{V_x\}_{x\in\mathcal{K}}$ is an open cover of $\mathcal{K}$ and invoking compactness again we can extract a finite subcover $\{V_{x_i}\}_{i=1,\ldots,m}$. At this point we let $g(z) = \max_i g_{x_i}(z) \in \bar{\mathcal{A}}$ and note that $g(z) > f(z)$ for all $z \in \mathcal{K}$ and since $g_x < f + \varepsilon$ for all x we get $f < g < f + \varepsilon$. □

Theorem 3. *If $\mathcal{K}$ is a compact metrizable space then $C(\mathcal{K})$ is separable.*

Proof. Take the functions $f_{x,n}(z) = (1 - n d(x, z))_+$. These functions are continuous, positive and $f_{x,n}(z) = 0$ if $d(x, z) > 1/n$, moreover the sets $V_{x,n} = \{z: f_{x,n}(z) > 0\}$ are open and $x \in V_{x,n}$ so for every $n \geqslant 1$ the family $\{V_{x,n}\}_{x\in\mathcal{K}}$ is an open cover of $\mathcal{K}$ and by compactness we can extract a finite subcover $\{V_{x_i^n,n}\}_{i=1,\ldots,N_n}$, by collecting all the subcovers we can form a countable set of functions $\{f_{x_i,n_i}\}_{i\geqslant 1}$ for which $n_i \to \infty$ when $i \to \infty$. This set separates the points in $\mathcal{K}$. Indeed if $x \neq y \in \mathcal{K}$ then $d(x, y) = c > 0$ and thus for i large enough there exists a function f_{x_i,n_i} such that $f_{x_i,n_i}(x) > 0$ and $f_{x_i,n_i}(y) = 0$. Taking all polynomials of these functions we get an algebra which is still countable and which separates the points, so by the Stone-Weierstrass theorem it is dense in $C(\mathcal{K})$. □

Remark 4. Separability can fail for different reasons. Note that metric spaces that are not separable cannot be compact.

1. In the large. If the ambient space is not compact separability needs not to hold in general. A basic example of non-separable space is the Banach space of all bounded sequences indexed by $\mathbb{N}$: $\ell^\infty = \{a = (a_n)_{n \geqslant 1} : \|a\| = \sup_n |a_n| < +\infty\}$. To see why, note that for any $A \subseteq \mathbb{N}$ we can set $a_n^A = 1_{n \in A}$ and then $a^A \in \ell^\infty$ but $\|a^A - a^B\| = 1 \Leftrightarrow A \neq B$ which means that the exists an uncountable infinity of points which are at distance 1. No countable set can be used to approximate all these points at the same time.

2. In the small. The space $L^\infty([0, 1], \mathrm{d}x)$ with the topology induced by the sup norm is not separable. Just observe that for any $\delta > 0$ and any $x \in [\delta, 1 - \delta]$ the functions $f_x(z) = 1_{z \in B_{x,\delta}}$ are such that $\|f_x - f_y\| = 1$ if $x \neq y$ and that they are uncountable. Like in ℓ^∞, the topology here is too sensible to the details.

3. For any $1 \leqslant p < \infty$ $L^p([0, 1], \mathrm{d}x)$ is separable. A probabilistic argument follows. Consider the probability space $[0, 1]$ with the Borel σ-algebra $\mathcal{F}$ and the σ-algebras $\mathcal{F}_n$ generated by dyadic intervals of size 2^{-n}. Observer that $\mathcal{F}_\infty = \mathcal{F}$. Then for any $\varphi \in L^p$ we can consider the martingale $\varphi_n = \mathbb{E}[\varphi | \mathcal{F}_n]$ and by the martingale convergence theorem we have that $\varphi_n \to \varphi$ in L^p. It is enough then to approximate $\mathcal{F}_n$ measurable real bounded functions by $\mathcal{F}_n$ measurable bounded functions with values in $\mathbb{Q}$ which is a countable subset of $L^p([0, 1], \mathrm{d}x)$.

We denote $\Pi(\mathcal{K})$ the set of all Borel probability measures on $\mathcal{K}$. Any $\mu \in \Pi(\mathcal{K})$ defines a linear functional on $C(\mathcal{K})$ by integration: $f \mapsto \int_{\mathcal{K}} f(x)\mu(\mathrm{d}x)$. By abuse of notation we will still denote by μ this functional, so $\mu(f) = \int_{\mathcal{K}} f(x)\mu(\mathrm{d}x)$. It is a positive linear functional ($f \geqslant 0 \Rightarrow \mu(f) \geqslant 0$) and moreover $\mu(1) = 1$. Actually there is a one-to-one correspondence between such functionals and Borel probability measures on $\mathcal{K}$.

Theorem 5. (Riesz-Markov) *For any positive linear functional ℓ on the space $C(\mathcal{K})$ there exists a unique Borel measure μ on $\mathcal{K}$ such that*

$$\ell(f) = \int_{\mathcal{K}} f(x)\mu(\mathrm{d}x).$$

For a proof see [Reed and Simon, Functional analysis, vol 1, Academic Press, Th IV.14 and IV.18].

Remark 6. In the Riesz-Markov theorem compactness is necessary. Consider the functional ℓ: $C_b(\mathbb{R}) \to \mathbb{R}$ ($C_b(\mathbb{R})$ is the space of bounded continuous functions) defined as $\ell(f) = \lim_{x \to +\infty} f(x)$ when the limit exists and extended by the Hanh-Banach theorem to a linear functional on the whole $C_b(\mathbb{R})$. It is clear that a measure μ representing ℓ does not exist, in some sense ℓ is concentrated “at infinity”.

A sequence of elements $\mu_n \in \Pi(\mathcal{K})$ weakly converge to $\mu \in \Pi(\mathcal{K})$ if $\mu_n(f) \to \mu(f)$ for all $f \in C(\mathcal{K})$. Endowed with the topology associated to this convergence the space $\Pi(\mathcal{K})$ is metrizable, complete and separable, it is then a Polish space. A possible metric is determined by the countable dense set $\{f_n\}_{n \geqslant 1} \subseteq C(\mathcal{K})$ as follows

$$d(\mu, \nu) = \sum_{n \geqslant 1} \frac{|\mu(f_n) - \nu(f_n)|}{2^n \|f_n\|}.$$

Exercise 1. Verify that it is a metric. Prove that $d(\mu_n, \mu) \to 0 \Leftrightarrow \mu_n(f) \to \mu(f)$. Hint: $|\mu_n(f) - \mu(f)| \leqslant 2\|f_k - f\| + |\mu_n(f_k) - \mu(f_k)| \leqslant 2\|f_k - f\| + 2^k\|f_k\| d(\mu_n, \mu)$.

Some other remarks on semi-continuous functions. If A is an open set the function $1_A(x)$ is not continuous but it is lower semi-continuous.

Lemma 7. *A function $\varphi \colon \mathcal{K} \to \mathbb{R}$ is lower semi-continuous (lsc) iff the following equivalent statements hold:*

a) For all $x_k, x \in \mathcal{K}$ $\liminf_k \varphi(x_k) \geqslant \varphi(x)$.

b) For all $c \in \mathbb{R}$ the set $\{x \colon \varphi(x) \leqslant c\}$ is closed.

c) The function φ is the supremum of a family of continuous functions.

d) There exists $f_n \in C(\mathcal{K})$ such that $f_n(x) \uparrow \varphi(x)$ for all $x \in \mathcal{K}$.

e) For each x, $\lim_{\varepsilon \to 0} \inf_{y \in B_{x,\varepsilon}} \varphi(y) = \varphi(x)$ where $B_{x,\varepsilon} = \{y \colon d(x, y) \leqslant \varepsilon\}$.

Proof. Exercice. (d) $\Rightarrow$ (c) trivial; (c) $\Rightarrow$ (b) easy; (b) $\Rightarrow$ (a) consider $\{y \in \mathcal{K} \colon f(y) \leqslant f(x) - \varepsilon\}$; (a) $\Rightarrow$ (d) consider $f_n(x) = \inf_{y \in \mathcal{K}} (\varphi(y) + n\, d(x, y))$. □

Remark 8. Analogous properties hold for upper semi-continuous functions (usc) φ (which are such that $-\varphi$ is lower semi-continuous).

Exercise 2. Show that a lsc function attains its minimum (Prove it using compactness). Show that if g is lsc and f is non-decreasing (and continuous) then $f \circ g$ is lsc.