STONE-WEIERSTRASS THEOREM-Notes.

Stone-Weierstrass approximation theorem.

Let A be a vector space over R. A is an *algebra* (or: commutative algebra with unit) if there exists a 'multiplication operation' $A \times A \to A$, $(f,g) \mapsto f*g$ which is bilinear (linear in f and g), commutative ($f*g = g*f$) and there is an element $1 \in A$ (the 'unit') satisfying $f*1 = 1*f = f$ for all $f \in A$. A *subalgebra* of A is a vector subspace of A which is closed under the multiplication operation and contains the unit.

For any metric space X, the space $C^b_R(X)$ of continuous, bounded real-valued functions on X is an *algebra* over the field of real numbers (with the operation of pointwise multiplication), satisfying, for the uniform norm:

$$||fg|| \leq ||f||||g||.$$

The unit is the constant function 1.

Exercise 1. Use this to show that the closure $\bar{A}$ of any subalgebra of $C^b_R(X)$ is also a subalgebra. *Hint:* recall $\bar{A}$ is the set of functions in $C^b_R(X)$ which are uniform limits of functions in A. Given $f = \lim f_n, g = \lim g_n$, the main point is checking that $fg = \lim(f_n g_n)$. Estimate $|fg - f_n g_n|(x)$ in the natural way.

Definition. We say a sub algebra of $C^b_R(X)$ *separates points* if $\forall x \neq y$ in X we may find $f \in A$ with $f(x) \neq f(y)$.

Example. The polynomial functions in one variable (with real coefficients) form a subalgebra of $C^b_R([0,1])$. The polynomial functions in n real variables form a subalgebra of $C^b_R([0,1]^n)$. Both of these separate points.

The polynomials in one variable made up of even-degree monomials also form a subalgebra of $C^b_R([-1,1])$, which doesn't separate points (any such polynomial takes the same value at 1 and -1).

Interpolation property. Assume $A \subset C^b_R(X)$ is a subalgebra that separates points. For all $x \neq y$ in X and all real numbers a, b, there exists $f \in A$ so that $f(x) = a, f(y) = b$.

By assumption we know there exists $g \in A$ so that $g(x) \neq g(y)$. Set:

$$f = a + (b-a)\frac{g - g(x)}{g(y) - g(x)}.$$

(Note that adding a constant to an element of A yields another element of A, since the unit (the constant function 1) is in A.)

Stone-Weierstrass theorem. Let X be a compact metric space, $A \subset C_R(X)$ a subalgebra containing the constants and separating points. Then A is dense in the Banach space $C_R(X)$.

Main Lemma. The pointwise max and the pointwise min of finitely many functions in $\bar{A}$ is still in $\bar{A}$.

We first give the proof of the theorem assuming the main lemma, then prove the lemma. There are two steps:

Step 1. Given $f \in C_R(X), x \in X$ and $\epsilon > 0$, we find $g_x \in \bar{A}$ so that $g(x) = f(x)$ and $g_x(y) \leq f(y) + \epsilon$, for all $y \in X$.

Step 2. Using compactness, argue there are finitely many points $x_1, \ldots, x_N \in X$ so that $\varphi(x) = \max\{g_{x_1}(x), \ldots, g_{x_N}(x)\}$ (which is in $\bar{A}$, by the main lemma) satisfies:

$$f(y) - \epsilon \leq \varphi(y) \leq f(y) + \epsilon,$$

for all $y \in X$. Thus for any $\epsilon > 0$ we may find $\varphi \in \bar{A}$ so that $||f - \varphi|| \leq \epsilon$ (uniform norm). So $f \in \bar{A}$.

Step 1. For each $f \in C_R(X)$, each $x \in X$ and any $\epsilon > 0$, there exists a function $g \in \bar{A}$ so that $g(x) = f(x)$ and $g(y) \leq f(y) + \epsilon \forall y \in X$.

Proof. Given $z \in X$ with $z \neq x$, let $h_z \in \bar{A}$ satisfy $h_z(x) = f(x)$ and $h_z(z) = f(z) + \epsilon/2$ (from the interpolation property.) By continuity, there is an open neighborhood V_z of z in X so that, for each $y \in V_z$ we have $h_z(y) \leq f(y) + \epsilon$. These define an open cover $\{V_z\}_{z \in X}$ of X. Taking a finite subcover $\{V_{z_i}\}_{i=1}^N$ of X, we find (from the Main Lemma) the function $g = \min\{h_{z_i} | i = 1, \ldots, N\}$ is in $\bar{A}$ and satisfies the conditions required.

Proof of Step 2.

Let $f \in C_R(X)$ be arbitrary Given $\epsilon > 0$ and $x \in X$, let $g_x \in \bar{A}$ be the function from Step 1. By continuity there is a neighborhood $U(x)$ of x in X so that $g_x(y) \geq f(y) - \epsilon$ for $y \in U(x)$. Cover X by a finite number of neighborhoods $U(x_i)$, $i = 1, \ldots, N$. Then (from the Main Lemma) the function $\varphi = \max(g_{x_i})$ is in $\bar{A}$ and satisfies $f(y) - \epsilon \leq \varphi(y) \leq f(y) + \epsilon$.

Prior to proving the Main Lemma, we need a result of general interest. It is easy to give examples of sequences of continuous functions converging *non-uniformly* to a continuous function. (For example, consider $x^n(1 - x^n)$ in $[0, 1]$.) However, this can't happen for monotone sequences on compact spaces:

Dini's theorem. Let X be a compact metric space. If an increasing (or decreasing) sequence (f_n) of continuous real-valued functions on X converges pointwise to a continuous function f, then the convergence is uniform.

Proof. Given $\epsilon > 0$ and $x \in X$ we may find an integer $n(x)$ so that $0 \leq f(x) - f_{n(x)}(x) \leq \epsilon$. By continuity (of f and $f_{n(x)}$ at x), we may find a neighborhood $V(x)$ of x in X so that:

$$|f(x) - f(y)| \leq \epsilon \text{ and } |f_{n(x)}(x) - f_{n(x)}(y)| \leq \epsilon, \text{ for all } y \in V(x).$$

Then for each $y \in V(x)$ we have $0 \leq f(y) - f_{n(x)}(y) \leq 3\epsilon$. Take a finite subcover of $\{V(x)\}_{x \in X}$ and the maximum N of the $n(x_i)$. Then for each $n \geq N$ we have $f(y) - f_n(y) \leq f(y) - f_{n(x_i)}(y) \leq 3\epsilon$, if $y \in V(x_i)$. Since the $V(x_i)$ cover X, this ends the proof.

Question. What goes wrong in this proof if the sequence is not monotone?

Proof of the Main Lemma.

Step 1. There exists a sequence (u_n) of real polynomials approximating $\sqrt{t}$ uniformly in $[0, 1]$.

Define u_n by recurrence, letting $u_1 = 0$ and setting:

$$u_{n+1}(t) = u_n(t) + \frac{1}{2}(t - u_n(t)^2).$$

We show by induction that $u_{n+1} \geq u_n$ and $u_n(t) \leq \sqrt{t}$ in $[0, 1]$. It follows from the recursion relation that the first fact follows from the second. On the other hand,

$$\sqrt{t} - u_{n+1}(t) = \sqrt{t} - u_n(t) - \frac{1}{2}(t - u_n^2(t)) = (\sqrt{t} - u_n(t))(1 - \frac{1}{2}(\sqrt{t} + u_n(t))$$

and from $u_n(t) \leq \sqrt{t}$ it follows that the second factor is positive. Thus we have pointwise convergence of u_n to $\sqrt{t}$ (from the recurrence relation), and then uniform convergence follows from Dini's theorem.

Exercise 2. Compute the approximations $u_n(t)$ for $n = 1, 2, 3, 4$, and plot them in $[0, 1]$ (on the same graph).

Step 2. If $f \in \bar{A}$, then $|f| \in \bar{A}$, the closure of A in $C_R(X)$.

Let $a = ||f||$ (sup norm). The function f^2/a^2 is in $\bar{A}$ (since $\bar{A}$ is an algebra) and takes values in $[0, 1]$. If $u_n(t)$ are the functions from Step 1, the compositions $u_n \circ (f^2/a^2)$ are in $\bar{A}$ (again since A is an algebra), and converge uniformly to $|f|/a$.

Step 3. If $f, g \in A$, then $\max\{f, g\}, \min\{f, g\}$ are also in $\bar{A}$.

$$\max\{f, g\} = \frac{1}{2}(f + g + |f - g|), \quad \min\{f, g\} = \frac{1}{2}(f + g - |f - g|).$$

Step 4 (final). The pointwise max and the pointwise min of finitely many functions in $\bar{A}$ is still in $\bar{A}$: follows from Step 3, since minimizing over a finite set amounts to a finite number of pair comparisons.

Remark. The theorem is false for subalgebras of $C_{\mathbb{C}}(X)$ (complex-valued functions). This follows from the classical result in Complex Analysis:

Let $f_n : D \to \mathbb{C}$ be a sequence of complex analytic functions in a domain $D \subset \mathbb{C}$. Suppose $f_n \to f$ uniformly on compact subsets of D. Then $f : D \to \mathbb{C}$ is analytic in D.

Corollary 1. (Polynomials in R^n.) Any real-valued continuous function on a compact subset of R^n is the uniform limit of a sequence of polynomials.

Corollary 2. (Separability.) If X is a compact metric space, the space $C_R(X)$ is separable.

First note that *a compact metric space X is separable.* Indeed covering X by open balls of radius $1/n$ we see that for each $n \geq 1$ there is a finite set A_n so that, for each $x \in X$, $d(x, A_n) \leq 1/n$. Then $A = \cup_n A_n$ is countable, and it is easy to see that $\bar{A} = X$.

Exercise 6. Prove this: for any $x \in X$, there exists a sequence $(x_j)_{j \geq 1}$ of points of A, so that $d(x_j, x) \to 0$.

In general, *any separable metric space is second-countable* (that is, has a countable basis of open sets): Let D be a countable dense set, $\mathcal{B}$ the countable set of open balls with rational radius and center a point of D. Let U be open, and $x \in U$. Then there is a ball $B_r(x) \subset U$ with r rational, and we can find some $d \in D \cap B_{r/3}(x)$. Then $x \in B_{2r/3}(d) \subset B_r(x) \subset U$, and $B_{2r/3}(d)$ is in $\mathcal{B}$.

Let (U_n) be a countable basis for the topology of X, and let $g_n(x) = d(x; X \setminus U_n)$. The monomials $g_1^{m_1} \dots g_r^{m_r}$ (with the m_j integers) form a countable set (h_n) of continuous functions on X, and the vector space they span is the algebra A generated by the g_n. So it suffices to use the Stone-Weierstrass theorem to conclude A is dense in $C_R(X)$.

The family $\{g_n\}$ separates points: if $x \neq y$, we may find an U_n so that $x \in U_n, y \notin U_n$, and thus $g_n(x) \neq 0, g_n(y) = 0$.

Alternative approach: Bernstein polynomials.

It is a remarkable fact that, for uniform approximation by polynomials in the unit interval $[0, 1]$, there is an explicit procedure that amounts almost to “a formula”.

Denote by $C_j^n = C_{n-j}^n$ the binomial coefficient: $C_j^n = \frac{n!}{j!(n-j)!}, 0 \leq j \leq n$. As we learn in high school:

$$\sum_{j=0}^{n} C_j^b x^j (1-x)^{n-j} = (x+1-x)^n = 1, \quad x \in [0,1].$$

We use these terms as coefficients and, for each $n \geq 1$, 'sample' the function $f \in C[0,1]$ at equidistant points to define the polynomial $B_n[f](x)$:

$$B_n[f](x) = \sum_{j=0}^{n} f(\frac{j}{n}) C_j^n x^j (1-x)^{n-j}.$$

Theorem: $B_n[f] \to f$ uniformly in $[0,1]$.

Proof. First note that $B_n[f](0) = f(0), B_n[f](1) = f(1)$. Then, letting $q_{nj}(x) = C_j^n x^j (1-x)^{n-j}$, we have:

$$\sum_{j=0}^{n} q_{nj}(x) \equiv 1 \Rightarrow |f(x) - B_n[f](x)| \leq \sum_{j=0}^{n} |f(x) - f(\frac{j}{n})| q_{nj}(x).$$

By uniform continuity of f, given $\epsilon > 0$ we may find $\delta > 0$ (depending only on ϵ and f) so that $|f(x) - f(\frac{j}{n})| < \epsilon$ whenever $|x - \frac{j}{n}| < \delta$. So for each $x \in [0,1]$ we split the points $\frac{j}{n}$ in $[0,1]$ into two sets:

$$N_1 = \{j = 1, \ldots, n; |x - \frac{j}{n}| < \delta\}, \quad N_2 = \{j = 1, \ldots, n; |x - \frac{j}{n}| \geq \delta\}.$$

The sum over N_1 is easy to estimate:

$$\sum_{j \in N_1} |f(x) - f(\frac{j}{n})(x)| q_{nj}(x) < \epsilon \sum_{j=0}^{n} q_{nj}(x) = \epsilon.$$

To estimate the other sum, we need a lemma.

Lemma.$\sum_{j=0}^{n} q_{nj}(x)(x - \frac{j}{n})^2 = \frac{x(1-x)}{n} \leq \frac{1}{4n}$.

Assuming the lemma, with $|f(x)| \leq M$ in $[0,1]$ we have:

$$\sum_{j \in N_2} |f(x) - f(\frac{j}{n})(x)| q_{nj}(x) \leq 2M \sum_{j=0}^{n} q_{nj}(x) \frac{(x - \frac{j}{n})^2}{\delta^2} \leq \frac{M}{2n\delta^2} < \epsilon,$$

provided $n > M/2\delta^2$. This concludes the proof.

Proof of Lemma. Expanding $(x - \frac{j}{n})^2$, we see it is enough to compute:

$$B_n[1](x) = \sum_{j=0}^{n-1} q_{nj}(x) = 1;$$

using $(j/n)C_j^n = C_{j-1}^{n-1}$:

$$B_n[x](x) = \sum_{j=0}^{n} q_{nj}(x)\frac{j}{n} = x\sum_{j=1}^{n} C_{j-1}^{n-1}x^{j-1}(1-x)^{(n-1)-(j-1)} = x\sum_{k=0}^{n-1} q_{(n-1)k}(x) = x.$$

$$B_n[x^2](x) = \sum_{j=0}^{n}(\frac{j}{n}C_j^n)x^j(1-x)^{n-j}(\frac{j}{n}) = \frac{x}{n}\sum_{j=1}^{n}(j-1)C_{j-1}^{n-1}x^{j-1}(1-x)^{(n-1)-(j-1)}+\frac{x}{n}$$

$$= \frac{x^2}{n}(n-1)\sum_{j=2}^{n} C_{j-2}^{n-2}x^{j-2}(1-x)^{n-2-(j-2)} + \frac{x}{n} = x^2 + \frac{1}{n}x(1-x).$$

Remark 1. Note that this computes the Bernstein polynomials of $1, x, x^2$. In particular, 1 and x are eigenfunctions of the linear operator B_n in $C[0,1]$, with eigenvalue 1.

Exercise 3. Do the calculation that completes the proof of the Lemma.

Exercise 4. Show that $f(x) = x(1-x)$ is an eigenfunction of the linear operator B_n in $C[0,1]$, with eigenvalue $\lambda = (n-1)/n$. (This means $B_n[f] = \lambda f$.)

Exercise 5. Compute $B_n[f](x)$ for $f(x) = \sqrt{x}$ and $n = 1, 2, 3, 4$, and plot them in $[0,1]$.

Remark 2. The following is sometimes called *Fundamental Theorem on Approximation in Normed Vector spaces:* If V is a finite-dimensional vector space of a normed vector space E, then for every $f \in E$ there exists at least one best approximation $p \in V$. (For example, $E = C_R^b([0,1])$, V the subspace of real-valued polynomials of degree n, restricted to $[0,1]$).

The theorem follows from the fact that V is closed in E, so if $f \in E \setminus V$ the distance $d(f,V) = \inf\{||f-p||; p \in V\}$ is positive, and attained by some vector $p \in V$. (By a previous exercise.)

Remark 3. We have the following *quantitative error estimate* for approximation by $B_n[f]$ in $C_R[0,1]$:

$$||f - B_n[f]|| \leq \frac{5}{4}\omega_f(\frac{1}{\sqrt{n}}).$$

Here ω_f is the *modulus of continuity* of the continuous function f:

$$\omega_f(\delta) = \sup\{|f(x) - f(y)|; |x - y| \leq \delta, x, y \in [0, 1]\}.$$

For example, if f is Hölder continuous with exponent $\alpha \in (0, 1)$: $\omega_f(\delta) \leq K\delta^\alpha$.