MATH31001/MATH41001/MATH61001: LINEAR ANALYSIS

CHAPTER 1: APPROXIMATING CONTINUOUS FUNCTIONS

NOTATION AND BASIC FACTS

Notation. We shall use $\mathbb{R}$ to denote the set of real numbers and $\mathbb{C}$ to denote the set of complex numbers.

The modulus or absolute value on $\mathbb{R}$ is denoted by $|x|$:

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

and the distance between two points $x, y \in \mathbb{R}$ is given by $|x - y|$.

The modulus on $\mathbb{C}$ is denoted by $|z|$. If $z = x + iy$ (with $x, y \in \mathbb{R}$) then

$$|z| = (x^2 + y^2)^{1/2}.$$

(Note that, if $z \in \mathbb{R}$, this agrees with the definition above.) Again, we may define the distance between $z, w \in \mathbb{C}$ by $|z - w|$.

For $n \geq 1$, we shall let $\mathbb{R}^n$ denote the n-dimensional real space and $\mathbb{C}^n$ the n-dimensional complex space:

$$\mathbb{R}^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\}$$

and

$$\mathbb{C}^n = \{(z_1, \dots, z_n) : z_1, \dots, z_n \in \mathbb{C}\}.$$

Write $x = (x_1, \dots, x_n)$. The natural norm on $\mathbb{R}^n$ (or $\mathbb{C}^n$) is defined by

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}.$$

We shall sometimes denote this norm by $\| \cdot \|_2$ when we want to distinguish it from other norms; this will be made clear at the time.

We may use the norm to define a distance function (or metric) $d(\cdot, \cdot)$ on $\mathbb{R}^n$ or $\mathbb{C}^n$:

$$d(x, y) = \|x - y\| = \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2}.$$

Typeset by $\mathcal{AMS}$-TEX

Metric spaces. A metric space is a set X together with a distance function (or metric) $d : X \times X \to \mathbb{R}$ such that

(1) $d(x,x) = 0$ and $d(x,y) > 0$ if $x \neq y$;
(2) $d(x,y) = d(y,x)$;
(3) $d(x,y) \leq d(x,z) + d(z,y)$.

For $x \in X$ and $\epsilon > 0$, we define the ϵ-ball $B(x,\epsilon)$ centred at x by

$$B(x,\epsilon) = \{y \in X \, : \, d(x,y) < \epsilon\}.$$

Here are some basic definitions are properties of metric spaces.

- We say that $E \subset X$ is *dense* in X if, for every $x \in X$ and $\epsilon > 0$, there exists $y \in E$ such that $y \in B(x,\epsilon)$.
- We call a set $U \subset X$ an *open* set if, for every $x \in U$, there exists $\epsilon > 0$ such that $B(x,\epsilon) \subset U$. (Note that the empty set $\varnothing$ and X itself are open sets.)
- We call a set $C \subset X$ a *closed* set if C contains its limit points, i.e., if $\{x_n\}_{n=1}^\infty \subset C$ and $\lim_{n\to+\infty} x_n = x$ then $x \in C$.
- The *closure* of a set $E \subset X$, denoted by $\overline{E}$, is the union of E and all its limit points (or, equivalently, it is the smallest closed set containing E). The set E is dense in X if and only if $\overline{E} = X$.
- We say that a sequence $\{x_n\}_{n=1}^\infty$ in X is a *Cauchy sequence* if, for all $\epsilon > 0$, there exists $N \geq 1$ such that if $n, m \geq N$ then $d(x_n, x_m) < \epsilon$.
- We say that X is *complete* if every Cauchy sequence $\{x_n\}_{n=1}^\infty$ in X converges, i.e., there exists $x \in X$ such that $\lim_{n\to+\infty} d(x_n, x) = 0$. More generally, we say that $E \subset X$ is complete if every Cauchy sequence in E converges to a point in E.
- If X is complete and $C \subset X$ is closed then C is complete. (If $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence in C then, since X is complete, it converges to some $x \in X$. However, since C is closed, $x \in C$.)
- Let K be a subset of X. If $\{U_i\}_{i\in\mathcal{I}}$ is a collection of open sets such that $K \subset \bigcup_{i\in\mathcal{I}} U_i$ then we call $\{U_i\}_{i\in\mathcal{I}}$ an open cover of K. (Here $\mathcal{I}$ is possibly infinite index set.) We say that K is a *compact* set if any open cover of K, as above, has a finite subcover, i.e., there exists $\{U_1, \ldots, U_n\} \subset \{U_i\}_{i\in\mathcal{I}}$ such that $K \subset \bigcup_{j=1}^n U_j$.
- We say that K is *sequentially compact* if every sequence in K has a subsequence which converges in K. If K is a subset of a metric space then K is sequentially compact if and only if K is compact.
- If K is compact then K is complete. (We do not need to assume that X is complete.)
- Suppose that Y is another metric space. A function $f : X \to Y$ is continuous if, for every open set $U \subset Y$, $f^{-1}U \subset X$ is open. This is equivalent to the ϵ-δ definition: $f : X \to Y$ is continuous if, for every $x \in X$ and for every $\epsilon > 0$, there exists $\delta > 0$ such that, for $y \in X$,

  $$d_X(x,y) < \delta \quad \Longrightarrow \quad d_Y(f(x), f(y)) < \epsilon.$$

  We shall denote the set of continuous functions from X to Y by $C(X,Y)$.
- If X is compact then any continuous function $f : X \to \mathbb{R}$ (or $\mathbb{C}$) is bounded, i.e., there exists $M \geq 0$ such that $|f(x)| \leq M$, for all $x \in X$.

Let us consider again the vector space $C([0,1],\mathbb{R})$ of continuous functions $f:[0,1]\to\mathbb{R}$. A norm on this space is given by

$$\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|;$$

this is called the uniform norm or the supremum norm. (Since $[0,1]$ is compact, f is bounded and so this quantity is finite.) This allows us to define a metric on $C([0,1],\mathbb{R})$ by

$$d_\infty(f,g) = \|f-g\|_\infty.$$

(It is conventional just to use norm notation in this context, so in fact we will write $\|f-g\|_\infty$ and not $d_\infty(f,g)$.)

We are going to show that for any such f and any $\epsilon > 0$, we can find a polynomial $p(x) = a_0 + a_1x + \ldots + a_nx^n$ such that, for all $x\in[0,1]$, $|f(x)-p(x)|\le\epsilon$. (The degree n of the polynomial is not fixed; we can take it as large as we need to get the approximation.)

In terms of the uniform norm $\|\cdot\|_\infty$ introduced above, the approximation condition may be written as $\|f-p\|_\infty\le\epsilon$.

Another way of phrasing the result is that polynomial functions are *uniformly dense* in $C([0,1],\mathbb{R})$, i.e., dense with respect to the metric defined by the uniform norm $\|\cdot\|_\infty$.

Let us state the result formally as a theorem.

Theorem 1.1 (Weierstrass Approximation Theorem). *Suppose that $f\in C([0,1],\mathbb{R})$ and that $\epsilon>0$. Then there exists a polynomial $p(x)$ such that $\|f-p\|_\infty\le\epsilon$.*

In fact, we shall show that f may be approxmated by polynomials of a particular form:

$$B_n(f;x) := \sum_{k=0}^{n}\binom{n}{k} f\left(\frac{k}{n}\right) x^k(1-x)^{n-k}.$$

This is called the *nth Bernstein polynomial for f*. Clearly it is a polynomial of degree n.

Before we prove the theorem, we need a lemma.

Lemma 1.2.

(i)

$$\sum_{k=0}^{n}\binom{n}{k} x^k(1-x)^{n-k} = 1.$$

(ii)

$$\sum_{k=0}^{n}k\binom{n}{k} x^k(1-x)^{n-k} = nx.$$

(iii)

$$\sum_{k=0}^{n}(k-nx)^2\binom{n}{k} x^k(1-x)^{n-k} = nx(1-x).$$

Proof.
(i) This follows from the Binomial Theorem.

(ii) Note that

$$\begin{aligned}\frac{d}{dx}\left(x^k(1-x)^{n-k}\right) &= kx^{k-1}(1-x)^{n-k} - (n-k)x^k(1-x)^{n-k-1}\\ &= x^k(1-x)^{n-k}\frac{k-nx}{x(1-x)}.\end{aligned}$$

Thus, differentiating (i) gives

$$\begin{aligned}0 &= \frac{d}{dx}\left(\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}\right)\\ &= \sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}\frac{k-nx}{x(1-x)}\end{aligned}$$

and, taking out the factor $1/x(1-x)$,

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}(k-nx) = 0. \tag{$*$}$$

Rearranging and using (i),

$$\sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k} = nx\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = nx,$$

as required.

(iii) Now differentiate (ii) to get

$$\begin{aligned}n &= \frac{d}{dx}\left(\sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k}\right)\\ &= \sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k}\frac{k-nx}{x(1-x)}.\end{aligned}$$

Thus

$$\sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k}(k-nx) = nx(1-x). \tag{$**$}$$

Note that

$$(k-nx)^2 = k(k-nx) - nx(k-nx).$$

Hence, using $(*)$ and $(**)$,

$$\begin{aligned}
&\sum_{k=0}^{n}(k-nx)^2\binom{n}{k}x^k(1-x)^{n-k} \\
&= \sum_{k=0}^{n}k(k-nx)\binom{n}{k}x^k(1-x)^{n-k} - nx\sum_{k=0}^{n}(k-nx)\binom{n}{k}x^k(1-x)^{n-k} \\
&= nx(1-x)+0 = nx(1-x),
\end{aligned}$$

as required. $\square$

Proof of Theorem 1.1. We shall now prove the Weierstrass Approximation Theorem. Fix $\epsilon > 0$. By uniform continuity of f, there exists $\delta > 0$ such that, for $x, y \in [0,1]$,

$$|x-y| < \delta \quad \Longrightarrow \quad |f(x)-f(y)| < \frac{\epsilon}{2}.$$

Using Lemma 1.2(i), we have

$$f(x) - B_n(f;x) = \sum_{k=0}^{n}\left(f(x) - f\left(\frac{k}{n}\right)\right)\binom{n}{k}x^k(1-x)^{n-k}.$$

Thus

$$|f(x)-B_n(f;x)| \le \sum_{k=0}^{n}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k} = \Sigma_1(x)+\Sigma_2(x),$$

where

$$\Sigma_1(x) = \sum_{\substack{0\le k\le n \\ k\,:\,\left|x-\frac{k}{n}\right|<\delta}}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k} < \frac{\epsilon}{2}$$

and

$$\begin{aligned}
\Sigma_2(x) &= \sum_{\substack{0\le k\le n \\ k\,:\,\left|x-\frac{k}{n}\right|\ge\delta}}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k} \\
&\le 2\|f\|_\infty \sum_{k\,:\,(k-nx)^2\ge n^2\delta^2}\binom{n}{k}x^k(1-x)^{n-k} \\
&\le 2\|f\|_\infty \frac{1}{n^2\delta^2}\sum_{k=0}^{n}(k-nx)^2\binom{n}{k}x^k(1-x)^{n-k} \\
&= 2\|f\|_\infty \frac{nx(1-x)}{n^2\delta^2} \\
&\le \frac{\|f\|_\infty}{2\delta^2 n},
\end{aligned}$$

where we have used Lemma 1.2(iii) (and the easy inequality $x(1-x) \leq 1/4$).

Combining the estimates on $\Sigma_1(x)$ and $\Sigma_2(x)$, we obtain

$$|f(x) - B_n(f;x)| \leq \frac{\epsilon}{2} + \frac{\|f\|_\infty}{2\delta^2 n}.$$

Now choose N sufficiently large that

$$\frac{\|f\|_\infty}{2\delta^2 N} < \frac{\epsilon}{2}.$$

(One may take $N = [\|f\|_\infty/\epsilon\delta^2] + 1$.) Then

$$|f(x) - B_N(f;x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

so $B_N(f;x)$ is a polynomial satisfying the conclusion of the theorem. $\square$

Remark. The rate of convergence of $B_n(f;\cdot)$ to f is very slow. In fact, if

$$\|f - B_n(f;\cdot)\|_\infty = o(n^{-1})$$

(i.e. $\lim_{n\to+\infty} n\|f - B_n(f;\cdot)\|_\infty = 0$) then f is linear ($f(x) = ax + b$).

Remark. Take $f(x) = e^x$ and $p_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$ (*not* a Bernstein polynomial). Then

$$|f(x) - p_n(x)| = \sum_{k=n+1}^\infty \frac{x^k}{k!} \leq \sum_{k=n+1}^\infty \frac{1}{k!} \leq \frac{1}{n!}.$$

This is incredibly fast *but* this function is real analytic, which is very rare.

THE STONE-WEIERSTRASS THEOREM

Theorem 1.1 is a special case of a much more general theorem valid for compact metric spaces. If X is a compact metric space, $C(X,\mathbb{R})$ will denote the set of continuous functions $f : X \to \mathbb{R}$. We can define the uniform norm on $C(X,\mathbb{R})$ by

$$\|f\|_\infty = \sup_{x\in X} |f(x)|.$$

Definition. We say that $\mathcal{A} \subset C(X,\mathbb{R})$ is an algebra if $\mathcal{A}$ is a linear subspace of $C(X,\mathbb{R})$ with the additional property that

$$f, g \in \mathcal{A} \quad \Longrightarrow \quad fg \in \mathcal{A}.$$

(Here fg is just the function obtained by pointwise multiplication: $(fg)(x) = f(x)g(x)$.)

Theorem 1.3 (Stone-Weierstrass Theorem). *Let X be a compact metric space. Let $\mathcal{A} \subset C(X,\mathbb{R})$ be an algebra such that*

1. *$\mathcal{A}$ contains a non-zero constant function;*
2. *$\mathcal{A}$ separates points (i.e., if $x, x' \in X$, $x \neq x'$, then there exists $f \in \mathcal{A}$ such that $f(x) \neq f(x')$.*

Then $\mathcal{A}$ is uniformly dense in $C(X,\mathbb{R})$.

Proof. Omitted. □

Remark. Since $\mathcal{A}$ is an algebra, if $\mathcal{A}$ contains a non-zero constant function then it contains all non-zero constant functions.

Example 1. Let $X = [0,1]$ and take $\mathcal{A}$ to be the set of polynomials on X. Then $\mathcal{A}$ contains the non-zero constant function 1 and the function $p(x) = x \in \mathcal{A}$ separates points, so the hypotheses of Theorem 1.3 are satisfied. This shows that the Weierstrass Approximation Theorem is a special case of the Stone-Weierstrass Theorem. However, the WAT is a component of the proof of the SWT, so the WAT needs to be proved independently.

Example 2. Let $X = [0,1]$ and

$$\mathcal{A} = \left\{ a_0 + \sum_{n=1}^{N} a_n \cos(2\pi n x) + \sum_{n=1}^{M} b_n \sin(2\pi n x) \,:\, a_n, b_n \in \mathbb{R}, N, M \geq 1 \right\}.$$

Then $\mathcal{A}$ satisfies hypothesis (1) (since $1 \in \mathcal{A}$) but not hypothesis (2) (since $f(0) = f(1)$ for all $f \in \mathcal{A}$). In fact, it is easy to see that $\mathcal{A}$ is *not* uniformly dense in $C([0,1],\mathbb{R}$: choose $g \in C([0,1],\mathbb{R}$ with $g(0) \neq g(1)$ and suppose that $2\epsilon < |g(0) - g(1)|$. Then, for any $f \in \mathcal{A}$, either $|g(0) - f(0)| \geq \epsilon$ or $|g(1) - f(1)| \geq \epsilon$, so that $\|g - f\|_\infty \geq \epsilon$.

However, the following simple modification of X allows us to make $\mathcal{A}$ a dense set.

Example 3. Let $X = [0,1]$ with 0 and 1 identified. (The parametrization $\{e^{2\pi i x} : 0 \leq x \leq 1\}$ identifies X with the unit circle. Take $\mathcal{A}$ as in Example 2. For $f \in \mathcal{A}$, the fact that $f(0) = f(1)$ now means that f is well-defined as a function on X. Furthermore, one can easily check that $\mathcal{A}$ now satisfies the hypotheses of Theorem 1.3. Therefore $\mathcal{A}$ is uniformly dense in $C(X,\mathbb{R})$.

We end the chapter with a complex version of the Stone-Weierstrass Theorem (for $C(X,\mathbb{C})$ instead of $C(X,\mathbb{R})$). Note that an additional hypothesis is required.

Theorem 1.4 (Complex Stone-Weierstrass Theorem). *Let X be a compact metric space. Let $\mathcal{A} \subset C(X,\mathbb{C})$ be an algebra such that*

1. *$\mathcal{A}$ contains a non-zero constant function;*
2. *$\mathcal{A}$ separates points (i.e., if $x, x' \in X$, $x \neq x'$, then there exists $f \in \mathcal{A}$ such that $f(x) \neq f(x')$;*
3. *if $f \in \mathcal{A}$ then its complex conjugate $\bar{f} \in \mathcal{A}$.*

Then $\mathcal{A}$ is uniformly dense in $C(X,\mathbb{R})$.

Proof. Omitted. □