MMAT5030 Notes 4

1 Inner Product Space and Orthogonal Projection

In the previous lectures we studied the uniform convergence of Fourier series. Since the limit of a uniformly convergent series of continuous functions is again continuous, the uniform convergence result only applies to continuous, piecewise 2π-periodic functions. When the function is piecewise smooth 2π-periodic, we have pointwise convergence at points of continuity and average convergence at jump discontinuity points. In this section we will measure the distance between functions by a norm weaker than the uniform norm. Under the new L^2-distance, you will see that every integrable function is equal to its Fourier expansion in a sense that will be described below.

Recall that there is an inner product defined on the n-dimensional Euclidean space called the Euclidean metric

$$< x, y >= \sum_{j=1}^{n} x_j y_j, \qquad x, y \in \mathbb{R}^n.$$

With this inner product, one can define the concept of orthogonality and angle between two vectors. Likewise, we can also introduce a similar product on the space of integrable functions. Specifically, for $f, g \in R[-\pi, \pi]$, the L^2**-product** is given by

$$< f, g >= \int_{-\pi}^{\pi} f(x)g(x)\, dx.$$

The L^2-product behaves like the Euclidean metric on $\mathbb{R}^n$ except at one aspect, namely, the condition $< f, f >= 0$ does not imply $f \equiv 0$. This is easy to see. In fact, when f is equal to zero except at finitely many points, then $< f, f >= 0$. It can be shown that $< f, f >= 0$ if and only if f is equal to zero except on a set of "measure zero" (see the end of this section). This minor difference with the Euclidean inner product will not affect our discussion much, except more caution is needed when we proceed. Parallel to the Euclidean case, we define the L^2-norm of an integrable function f to be

$$\|f\| = \sqrt{< f, f >},$$

and the L^2**-distance** between two integrable functions f and g by $\|f - g\|_2$. (When f, g are complex-valued, one should define the L^2-product to be

$$< f, g >= \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx \ ,$$

so that $< f, f >\geq 0$.) One can verify that the triangle inequality

$$\|f + g\| \leq \|f\| + \|g\|$$

holds. In fact, by taking square of both sides, this inequality reduces to showing

$$\text{Re} \int_{-\pi}^{\pi} f(x)\overline{g}(x)\, dx \leq \|f\|\|g\| \ ,$$

which is the Cauchy-Schwarz inequality.

We can also talk about $f_n \to f$ in L^2-sense, i.e., $\|f_n - f\| \to 0$, or equivalently,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |f_n(x) - f(x)|^2\, dx = 0, \qquad \text{as } n \to \infty \ .$$

This is a convergence in an average sense. It is not hard to see that when $\{f_n\}$ tends to f uniformly, $\{f_n\}$ must tend to f in L^2-sense. In fact, we have the inequality

$$\|f\|^2 \leq 2\pi \|f\|_\infty^2 ,$$

which means

$$\|f-g\| \leq \sqrt{2\pi}\|f-g\|_\infty ,$$

so uniform convergence is stronger than L^2-convergence. A moment's reflection will show that the converse is not always true. For instance, the functions $f_n(x) = 0, x \in [-\pi, -1/n] \cup [1/n, \pi]$, and $= n^{1/4}, x \in (-1/n, 1/n)$, satisfy $\|f_n - 0\| = \sqrt{2}n^{-1/4} \to 0$ but $\|f_n - 0\|_\infty \to \infty$ as $n \to \infty$. Hence convergence in L^2-sense is weaker than uniform convergence. It is interesting to observe that L^2-convergence and pointwise convergence are not compatible, see exercise.

Our aim is to show that the Fourier series of every integrable function converges to the function in the L^2-sense. In order to do this, it is necessary to study orthonormal sets in an inner product space. Let V be an inner product space and $u \in V$. The vector u is called normalized if $\|u\| = 1$ and two vectors u, v are **orthogonal** if $\langle u, v\rangle = 0$. A set is called **orthonormal** if all vectors in this set are normalized and pairwise orthogonal. It is not hard to show that all vectors in an orthonormal set must be linearly independent.

Let W be a subspace of V of dimension n and $S = \{w_1, w_2, \cdots, w_n\}$ be an orthonormal set in W. Since these vectors are also linearly independent, S forms an orthonormal basis of W. Now, we consider the following question: Given $u \in V$, how can we find the point $w^* \in W$ that minimizes the distance from u to W? In other words, we post

$$\min\{\|u-w\| : \ w \in W\}.$$

This problem can be posed without referring to an orthonormal basis. However, the key point is, using an orthonormal basis, there is a nice formula for w^*.

Theorem 4.1. Let $u \in V$ where V is an inner product space and W is a subspace spanned by the orthonormal set $\{w_1, w_2, \cdots, w_n\}$. The followings hold:

(a)

$$\|u-w^*\| \leq \|u-w\| \quad \forall w \in W ,$$

where $w^* = \sum_{j=1}^n \alpha_j w_j,\ \alpha_j = \langle u, w_j\rangle$, and equality holds if and only if $w = w^*$; and

(b)

$$\langle u-w^*, w\rangle = 0 , \quad \forall w \in W.$$

Proof. (a) To minimize $\|u-w\|$ is the same as to minimize $\|u-w\|^2$. Every w in W can be written as $w = \sum_{j=1}^n \beta_j w_j$, $\beta_j \in \mathbb{R}$. We have

$$\begin{aligned} \|u-w\|^2 &= \langle u-w, u-w\rangle \\ &= \langle u, u-w\rangle - \langle w, u-w\rangle \\ &= \|u\|^2 - 2\langle u, w\rangle + \|w\|^2 \\ &= \|u\|^2 - 2\sum_{j=1}^n \alpha_j\beta_j + \sum_{j=1}^n \beta_j^2. \end{aligned}$$

When $w = w^*$, we have $\alpha_j = \beta_j$, so

$$\|u-w^*\|^2 = \|u\|^2 - \sum_{j=1}^n \alpha_j^2 .$$

Therefore,

$$\|u - w^*\|^2 \le \|u - w\|^2$$

is the same as

$$\|u\|^2 - \sum_{j=1}^n \alpha_j^2 \le \|u\|^2 - 2\sum_{j=1}^n \beta_j \alpha_j + \sum_{j=1}^n \beta_j^2 \ .$$

But this follows readily from the inequality

$$\sum_{j=1}^n (\beta_j - \alpha_j)^2 \ge 0 \ .$$

It is also clear that the equality holds if and only if $\beta_j = \alpha_j$ for all j, that is, $w = w^*$. (a) is established.

Next, (b) follows from observing $\langle u - w^*, w_j \rangle = 0$ for all j.

The point $w^* \in W$ is called the **orthogonal projection** of u on W. Proposition 4.1(b) asserts that $u - w^*$ is orthogonal to W. In fact, this property uniquely characterizes w^*, see exercise. We point out $w^* = u$ if and only if $u \in W$.

2 Mean Convergence of Fourier Series

Now, we focus on the space $R[-\pi, \pi]$ with the L^2-product. Just like the canonical basis $\{e_1, \dots, e_n\}$ in $\mathbb{R}^n$, we knew that the functions

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx \right\}_{n=1}^\infty$$

forms an orthonormal set in $R[-\pi, \pi]$. (In the complex case, it is the set

$$\left\{ \frac{1}{\sqrt{2\pi}} e^{inx} \right\}_{n=-\infty}^\infty \ .$$

But we mainly consider the real case here.)

In the following we denote by

$$E_n = \left\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos jx, \frac{1}{\sqrt{\pi}} \sin jx \right\rangle_{j=1}^n$$

the $(2n+1)$-dimensional vector space spanned by the first $2n+1$ trigonometric functions.

First of all, taking $\{w_j\} = \{1/\sqrt{2\pi}, \cos jx/\sqrt{\pi}, \sin jx/\sqrt{\pi}\}$ and $W = E_n$ in Proposition 4.1, a direct computation shows that the orthogonal projection of f on E_n is precisely given by $S_n f$, where $S_n f$ is the n-th partial sum of the Fourier series of f. We can rewrite Proposition 4.1 in this special case as

Theorem 4.1' For $f \in R[-\pi, \pi]$ and for each $n \ge 1$,

(a)

$$\|f - S_n f\| \le \|f - g\| \ ,$$

and

(b)

$$\langle f - S_n f, g \rangle = 0 \ ,$$

for all g of the form

$$g = c_0 + \sum_{j=1}^{n} (c_j \cos jx + d_j \sin jx), \quad c_0, c_j, d_j \in \mathbb{R}.$$

Here is our main result.

Theorem 4.2. For every $f \in R[-\pi, \pi]$,

$$\lim_{n\to\infty} \|S_n f - f\| = 0.$$

Proof. Let $f \in R[-\pi, \pi]$. We further assume $f \geq 0$. For $\varepsilon > 0$, we can find a step function $s \geq 0$ such that $s \leq f$ and $\int_{-\pi}^{\pi} (f - s) < \varepsilon^2/16M$ where $M = \sup_x f(x)$. Then

$$\|f - s\| \leq \sqrt{M \int_{-\pi}^{\pi} (f - s)} = \frac{\varepsilon}{4} \ .$$

Next we modify s near its points of discontinuity to get a continuous, piecewise linear function g_1 satisfying

$$\|s - g_1\| < \frac{\varepsilon}{4}.$$

In case $g_1(\pi) \neq g_1(-\pi)$, we modify this function near π to get a new, piecewise linear function g satisfying $g(\pi) = g(-\pi)$ and

$$\|g - g_1\| < \frac{\varepsilon}{4} \ .$$

Now g is a continuous, piecewise linear (hence piecewise C^1-), 2π-periodic function. Appealing to Theorem 2.5 in Text, we can find some n_1 such that

$$\|g - S_n g\|_\infty < \frac{\varepsilon}{4\sqrt{2\pi}} \ , \quad \forall n \geq n_1 \ .$$

It implies

$$\|g - S_n g\| \leq \sqrt{2\pi} \|g - S_n g\|_\infty < \frac{\varepsilon}{4} \ .$$

Putting things together, we have, for all $n \geq n_1$,

$$\begin{aligned}
\|f - S_n f\| &\leq \|f - S_n g\| \quad \text{(by Proposition 4.1')} \\
&\leq \|f - s\| + \|s - g_1\| + \|g_1 - g\| + \|g - S_n g\| \\
&< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} \\
&= \varepsilon \ .
\end{aligned}$$

We have proved the theorem for non-negative functions. In the general case, we use the relation $f = f^+ - f^-$ and the triangle inequality to get

$$\begin{aligned}
\|f - S_n f\| &= \|f^+ - f^- - S_n f^+ + S_n f^-\| \\
&\leq \|f^+ - S_n f^+\| + \|f^- - S_n f^-\| \ .
\end{aligned}$$

Note that the use of Proposition 4.1' is the key to the proof in this theorem. As an application we have the following result concerning the uniqueness of the Fourier expansion.

Corollary 4.3. (a) Suppose that f_1 and f_2 in $R[-\pi,\pi]$ have the same Fourier series. Then f_1 and f_2 are equal almost everywhere.

(b) Suppose that f_1 and f_2 are continuous, 2π-periodic functions having the same Fourier series. Then f_1 is equal to f_2 everywhere.

Let $f = f_2 - f_1$. The Fourier coefficients of f all vanish, hence $S_n f = 0$, for all n. By Theorem 4.2, $\|f\| = \lim_{n\to\infty} \|f - S_n f\| = 0$. It follows that f^2, hence f, must vanish almost everywhere. In other words, f_2 is equal to f_1 almost everywhere. (a) holds. To prove (b), letting f be continuous and assuming that it is not equal to zero at some x_0, by continuity it is non-zero for all points near x_0. We can find some small $\delta > 0$ such that $f^2(x) \geq f^2(x_0)/2$ for all $x \in (x_0 - \delta, x_0 + \delta)$. But then

$$\begin{aligned} \int_{-\pi}^{\pi} f^2 &\geq \int_{x_0-\delta}^{x_0+\delta} f^2 \\ &\geq \frac{f^2(x_0)}{2} \times 2\delta > 0 \ , \end{aligned}$$

contradicting $\|f\| = 0$. Hence f must vanish identically.

Another interesting consequence of Theorem 4.2 is the Parseval's identity.

Corollary 4.4. (Parseval's Identity) For every $f \in R[-\pi,\pi]$,

$$\|f\|_2^2 = \frac{\pi}{2}a_0^2 + \pi \sum_{n=1}^{\infty} \left(a_n^2 + b_n^2\right),$$

where a_n and b_n are the Fourier coefficients of f.

Proof. Making use of the relations such as $\langle f, \cos nx/\sqrt{\pi}\rangle = \sqrt{\pi}a_n, n \geq 1$,

$$\begin{aligned} \langle f, S_n f\rangle &= \|S_n f\|_2^2 \quad \text{(by Proposition 4.1' (b))} \\ &= \frac{\pi}{2}a_0^2 + \pi \sum_{j=1}^{n} (a_j^2 + b_j^2). \end{aligned}$$

By Theorem 4.2 and $\langle f, S_n f\rangle = \|S_n f\|^2$ (Theorem 4.1'(b)),

$$\begin{aligned} 0 = \lim_{n\to\infty} \|f - S_n f\|^2 &= \lim_{n\to\infty} \left(\|f\|^2 - 2\langle f, S_n f\rangle + \|S_n f\|^2\right) \\ &= \lim_{n\to\infty} \left(\|f\|^2 - \|S_n f\|^2\right) \\ &= \|f\|^2 - \lim_{n\to\infty} \|S_n f\|^2 \\ &= \|f\|^2 - \left[\frac{\pi}{2}a_0^2 + \pi \sum_{n=1}^{\infty} \left(a_n^2 + b_n^2\right)\right]. \end{aligned}$$

In general, an orthonormal set $\{\phi_n\}$ in $R[a,b]$ is called **complete** if

$$\|f - \sum_{k=1}^{n} \langle f, \phi_k\rangle\|^2 \to 0, \qquad \text{as } n \to \infty \ ,$$

for every f. Whevever this happens, the proof above shows that the general Parseval's Identity

$$\|f\|_2 = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle^2$$

holds. Our main theorem asserts that $\{1/\sqrt{2\pi}, \cos nx/\sqrt{\pi}, \sin nx/\sqrt{\pi}\}$ forms a complete orthonormal set in $R[-\pi, \pi]$. It plays the role like the canonical basis $\{e_1, \cdots, c_n\}$ in the Euclidean space $\mathbb{R}^n$.

The norm of f can be regarded as the length of the "vector" f. Parseval's Identity shows that the square of the length of f is equal to the sum of the square of the length of the orthogonal projection of f onto each one-dimensional subspace spanned by the sine and cosine functions. This is an infinite dimensional version of the ancient Pythagoras theorem. It is curious to see what really comes out when you plug in some specific functions. For instance, we take $f(x) = x$ and recall that its Fourier series is given by $\sum 2(-1)^{n+1}/n \sin nx$. Therefore, $a_n = 0, n \geq 0$ and $b_n = 2(-1)^{n+1}/n$ and Parseval's identity yields Euler's summation formula

$$\frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots .$$

You could find more interesting identities by applying the same idea to other functions.

3 Appendix Sets of Measure Zero

This section is for reference only.

Let E be a subset of $\mathbb{R}$. It is called of measure zero, or sometimes called a null set, if for every $\varepsilon > 0$, there exists a (finite or infinite) sequence of intervals $\{I_k\}$ satisfying (1) $E \subset \cup_{k=1}^{\infty} I_k$ and (2) $\sum_{k=1}^{\infty} |I_k| < \varepsilon$. (When the intervals are finite, the upper limit of the summation should be changed accordingly.) Here I_k could be an open, closed or any other interval and its length $|I_k|$ is defined to the $b - a$ where $a \leq b$ are the endpoints of I_k.

The empty set is a set of measure zero from this definition. Every finite set is also null. For, let $E = \{x_1, \cdots, x_N\}$ be the set. For $\varepsilon > 0$, the intervals $I_k = (x_1 - \varepsilon/(4N), x_k + \varepsilon/(4N))$ clearly satisfy (1) and (2) in the definition.

Next we claim that every countable set is also of measure zero. Let $E = \{x_1, x_2, \cdots\}$ be a countable set. We choose

$$I_k = \left(x_k - \frac{\varepsilon}{2^{k+2}}, x_k + \frac{\varepsilon}{2^{k+2}} \right) .$$

Clearly, $E \subset \cup_{k=1}^{\infty} I_k$. On the other hand,

$$\begin{aligned} \sum_{k=1}^{\infty} |I_k| &= \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} \\ &= \frac{\varepsilon}{2} \\ &< \varepsilon . \end{aligned}$$

We conclude that every countable set is a null set.

There are uncountable sets of measure zero. For instance, the Cantor set which plays an important role in analysis, is of measure zero. Here we will not go into this.

The same trick in the above proof can be applied to the following situation.

Proposition 4.5. *The union of countably many null sets is a null set.*

Let $E_k, k \geq 1$, be sets of measure zero. For $\varepsilon > 0$, there are intervals satisfying $\{I_j^k\}, E_k \subset \cup_j I_j^k$, and $\sum_j |I_j^k| < \varepsilon/2^k$. It follows that $E \equiv \cup_k E_k \subset \cup_{j,k} I_j^k = \cup_k \cup_j I_j^k$ and

$$\sum_k \sum_j |I_j^k| < \sum_k \frac{\varepsilon}{2^k} = \varepsilon.$$

The concept of a null set comes up naturally in the theory of Riemann integration. A theorem of Lebsegue asserts that a bounded function is Riemann integrable if and only if its discontinuity set is null. (This result can be found in an appendix of Bartle-Sherbert and also in my 2060 notes. It will be proved again in Real Analysis. Presently you may simply take it for granted.) Let us prove the following result.

Proposition 4.6. *Let f be a non-negative integrable function on $[a,b]$. Then $\int_a^b f = 0$ if and only if f is equal to 0 except on a null set. Consequently, two integrable functions f, g satisfying*

$$\int_a^b |f(x) - g(x)|\, dx = 0,$$

if and only if f is equal to g except on a null set.

We set, for each $k \geq 1$, $A_k = \{x \in [a,b] : \ f(x) > 1/k\}$. It is clear that

$$\{x : \ f(x) > 0\} = \bigcup_{k=1}^{\infty} A_k \ .$$

By Proposition 4.5., it suffices to show that each A_k is null. Thus let us consider A_{k_0} for a fixed k_0. Recall from the definition of Riemann integral, for every $\varepsilon > 0$, there exists a partition $a = x_1 < x_2 < \cdots < x_n = b$ such that

$$0 \leq \sum_{k=1}^{n-1} f(z_k)|I_k| = \left| \sum_{k=1}^{n-1} f(z_k)|I_k| - \int_a^b f \right| < \frac{\varepsilon}{k_0} \ ,$$

where $I_k = [x_k, x_{k+1}]$ and z_k is an arbitrary tag in $[x_j, x_{j+1}]$. Let $\{k_1, \cdots, k_m\}$ be the index set for which I_{k_j} contains a point z_{k_j} from A_{k_0}. Choosing the tag point to be z_{k_j}, we have $f(z_{k_j}) > 1/k_0$. Therefore,

$$\frac{1}{k_0} \sum_{k_j} |I_{k_j}| = \sum_{k_j} f(z_{k_j})|I_{k_j}| \leq \sum_{k=1}^{n-1} f(z_k)|I_k| < \frac{\varepsilon}{k_0},$$

so

$$\sum_{k_j} |I_{k_j}| < \varepsilon.$$

We have shown that A_{k_0} is of measure zero.

Conversely, suppose f is equal to 0 except on a null set A. Since f is integrable, for any $\varepsilon > 0$, there is a partition $a = x_0 < x_1 < \cdots < x_n = b$ such that

$$\left| \int_a^b f - \sum_{k=1}^{n} f(z_k)(x_k - x_{k-1}) \right| < \varepsilon,$$

for any tag z_k in the interval $[x_{k-1}, x_k]$. Since A is a null set, for each k we can choose a tag $z_k \in [x_{k-1}, x_k] \cap A^c$, which yields $f(z_k) = 0$. Now

$$\begin{aligned}\left|\int_a^b f\right| &= \left|\int_a^b f - \sum_{k=1}^n f(z_k)(x_k - x_{k-1})\right| + \left|\sum_{k=1}^n f(z_k)(x_k - x_{k-1})\right| \\ &\le \varepsilon + \left|\sum_{k=1}^n (0)(x_k - x_{k-1})\right| \\ &= \varepsilon.\end{aligned}$$

Since ε is arbitrary, we have $\int_a^b f = 0$.

A property holds **almost everywhere** if it holds except on a null set. For instance, this proposition asserts that the integral of a non-negative function is equal to zero if and only if it vanishes almost everywhere.