Lecture 4

Inner product spaces

Of course, you are familiar with the idea of inner product spaces – at least finite-dimensional ones. Let $X$ be an abstract vector space with an inner product, denoted as $\langle \cdot, \cdot \rangle$, a mapping from $X \times X$ to $\mathbb{R}$ (or perhaps $\mathbb{C}$). The inner product satisfies the following conditions,

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$, $\forall x, y, z \in X$,
2. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$, $\forall x, y \in X$, $\alpha \in \mathbb{R}$,
3. $\langle x, y \rangle = \langle y, x \rangle$, $\forall x, y \in X$,
4. $\langle x, x \rangle \geq 0$, $\forall x \in X$, $\langle x, x \rangle = 0$ if and only if $x = 0$.

We then say that $(X, \langle \cdot, \cdot \rangle)$ is an inner product space.

In the case that the field of scalars is $\mathbb{C}$, then Property 3 above becomes

3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$, $\forall x, y \in X$,

where the bar indicates complex conjugation. Note that this implies that, from Property 2,

$$\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$$

$\forall x, y \in X$, $\alpha \in \mathbb{C}$.

Note: For anyone who has taken courses in Mathematical Physics, the above properties may be slightly different than what you have seen, as regards the complex conjugations. In Physics, the usual convention is to complex conjugate the first entry, i.e. $\langle \alpha x, y \rangle = \overline{\alpha} \langle x, y \rangle$.

The inner product defines a norm as follows,

$$\langle x, x \rangle = \|x\|^2 \quad \text{or} \quad \|x\| = \sqrt{\langle x, x \rangle}. \quad (1)$$

(Note: An inner product always generates a norm. But not the other way around, i.e., a norm is not always expressible in terms of an inner product. You may have seen this in earlier courses – the norm has to satisfy the so-called “parallelogram law.”)

And where there is a norm, there is a metric: The norm defined by the inner product $\langle \cdot, \cdot \rangle$ will define the following metric,

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}, \quad \forall x, y \in X. \quad (2)$$
And where there is a metric, we can discuss convergence of sequences, etc..

A complete inner product space is called a *Hilbert space*, in honour of the celebrated mathematician David Hilbert (1862-1943).

Finally, the inner product satisfies the following relation, called the “Cauchy-Schwarz inequality,”

\[ |\langle x, y \rangle| \leq \|x\|\|y\|, \quad \forall x, y \in X. \quad (3) \]

You probably saw this relation in your studies of finite-dimensional inner product spaces, e.g., \(\mathbb{R}^n\). It holds in abstract spaces as well.

**Examples:**

1. **\(X = \mathbb{R}^n\).** Here, for \(x = (x_1, x_2, \cdots, x_n)\) and \(y = (y_1, y_2, \cdots, y_n)\),

\[ \langle x, y \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \quad (4) \]

The norm induced by the inner product is the familiar Euclidean norm, i.e.

\[ \|x\| = \|x\|_2 = \left[ \sum_{i=1}^{n} x_i^2 \right]^{1/2}. \quad (5) \]

And associated with this norm is the Euclidean metric, i.e.,

\[ d(x, y) = \|x - y\| = \left[ \sum_{i=1}^{n} |x_i - y_i|^2 \right]^{1/2}. \quad (6) \]

You’ll note that the inner product generates only the \(p = 2\) norm or metric. \(\mathbb{R}^n\) is also complete, i.e., it is a Hilbert space.

2. **\(X = \mathbb{C}^n\) - a minor modification of the real vector space case.** Here, for \(x = (x_1, x_2, \cdots, x_n)\) and \(y = (y_1, y_2, \cdots, y_n)\),

\[ \langle x, y \rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}. \quad (7) \]

The norm induced by the inner product will be

\[ \|x\| = \|x\|_2 = \left[ \sum_{i=1}^{n} |x_i|^2 \right]^{1/2}, \quad (8) \]

and the associated metric is

\[ d(x, y) = \|x - y\| = \left[ \sum_{i=1}^{n} |x_i - y_i|^2 \right]^{1/2}. \quad (9) \]

\(\mathbb{C}^n\) is complete, therefore a Hilbert space.
3. The sequence space $l^2$ introduced earlier: Here, $x = (x_1, x_2, \cdots) \in l^2$ implies that $\sum_{i=1}^{\infty} x_i^2 < \infty$.

For $x, y \in l^2$, the inner product is defined as

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i.$$

(10)

Note that $l^2$ is the only $l^p$ space for which an inner product exists. It is a Hilbert space.

4. $X = C[a, b]$, the space of continuous functions on $[a, b]$ is NOT an inner product space.

5. The space of (real- or complex-valued) square-integrable functions $L^2[a, b]$ introduced earlier.

Here,

$$\|f\|^2 = \langle f, f \rangle = \int_{a}^{b} |f(x)|^2 \, dx < \infty.$$

(11)

The inner product in this space is given by

$$\langle f, g \rangle = \int_{a}^{b} f(x) \overline{g(x)} \, dx,$$

(12)

where we have allowed for complex-valued functions.

As in the case of sequence spaces, $L^2$ is the only $L^p$ space for which an inner product exists. It is also a Hilbert space.

6. The space of (real- or complex-valued) square-integrable functions $L^2(\mathbb{R})$ on the real line, also introduced earlier. Here,

$$\|f\|^2 = \langle f, f \rangle = \int_{-\infty}^{\infty} |f(x)|^2 \, dx < \infty.$$

(13)

The inner product in this space is given by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} \, dx,$$

(14)

Once again, $L^2$ is the only $L^p$ space for which an inner product exists. It is also a Hilbert space, and will the primary space in which we will be working for the remainder of this course.

**Orthogonality in inner product spaces**

An important property of inner product spaces is “orthogonality.” Let $X$ be an inner product space. If $\langle x, y \rangle = 0$ for two elements $x, y \in X$, then $x$ and $y$ are said to be orthogonal (to each other).

Mathematically, this relation is written as “$x \perp y$” (just as we do for vectors in $\mathbb{R}^n$).

We’re now going to need a few ideas and definitions for the discussion that is coming up.
• Recall that a **subspace** $Y$ of a vector space $X$ is a nonempty subset $Y \subset X$ such that for all $y_1, y_2 \in Y$, and all scalars $c_1, c_2$, the element $c_1 y_1 + c_2 y_2 \in Y$, i.e., $Y$ is itself a vector space. This implies that $Y$ must contain the zero element, i.e., $y = 0$.

• Moreover the subspace $Y$ is **convex**: For every $x, y \in Y$, the “segment” joining $x + y$, i.e., the set of all convex combinations,

$$ z = \alpha x + (1 - \alpha) y, \quad 0 \leq \alpha \leq 1, \quad (15) $$

is contained in $Y$.

• A vector space $X$ is said to be the **direct sum** of two subspaces, $Y$ and $Z$, written as follows,

$$ X = Y \oplus Z, \quad (16) $$

if each $x \in Z$ has a unique representation of the form

$$ x = y + z, \quad y \in Y, \; z \in Z. \quad (17) $$

The sets $Y$ and $Z$ are said to be **algebraic complements** of each other.

• In $\mathbb{R}^n$ and inner product spaces $X$ in general, it is convenient to consider spaces that are orthogonal to each other. Let $S \subset X$ be a subset of $X$ and define

$$ S^\perp = \{ x \in X \mid x \perp S \} = \{ x \in X \mid \langle x, y \rangle = 0 \; \forall y \in S \}. \quad (18) $$

The set $S^\perp$ is said to be the **orthogonal complement** of $S$.

**Note:** The concept of an algebraic complement does not have to invoke the use of orthogonality. With thanks to the student who raised the question of algebraic vs. orthogonal complementarity in class, let us consider the following example.

Let $X$ denote the (11-dimensional) vector space of polynomials of degree 10, i.e.,

$$ X = \{ \sum_{k=0}^{10} c_k x^k, \; c_k \in \mathbb{R}, \; 0 \leq k \leq 10 \}. \quad (19) $$

Equivalently,

$$ X = \text{span} \{ 1, x, x^2, \cdots, x^{10} \}. \quad (20) $$
Now define
\[ Y = \text{span} \{1, x, x^2, x^3, x^4, x^5\}, \quad Z = \text{span} \{x^6, x^7, x^8, x^9, x^{10}\}. \]  
(21)

First of all, \( Y \) and \( Z \) are subspaces of \( X \). Furthermore, \( X \) is a direct sum of the subspaces \( Y \) and \( Z \). However, the spaces \( Y \) and \( Z \) are not orthogonal complements of each other.

First of all, for the notion of orthogonal complementarity, we would have to define an interval of support, e.g., \([0,1]\), over which the inner products of the functions is defined. (And then we would have to make sure that all inner products involving these functions are defined.) Using the linearly independent functions, \( x^k, -\leq k \leq 10 \), one can then construct (via Gram-Schmidt orthogonalization) an orthogonal set of polynomial basis functions, \( \{\phi_k(x)\}, 0 \leq k \leq 10 \), over \( X \). It is possible that the first few members of this orthogonal set will contain the functions \( x^k, 0 \leq k \leq 5 \), which come from the set \( Y \). But the remaining members of the orthogonal set will contain higher powers of \( x \), i.e., \( x^k, 6 \leq k \leq 10 \), as well as lower powers of \( x \), i.e., \( x^k \mid 0 \leq k \leq 5 \). In other words, the remaining members of the orthogonal set will not be elements of the set \( Y \) – they will have nonzero components in \( X \).

See also Example 3 below.

**Examples:**

1. Let \( X \) be the Hilbert space \( \mathbb{R}^3 \) and \( S \subset X \) defined as \( S = \text{span}\{(1,0,0)\} \). Then \( S^\perp = \text{span}\{(0,1,0),(0,0,1)\} \). In this case both \( S \) and \( S^\perp \) are subspaces.

2. As before \( X = \mathbb{R}^3 \) but with \( S = \{(c,0,0) \mid c \in [0,1]\} \). Now, \( S \) is no longer a subspace but simply a subset of \( X \). Nevertheless \( S^\perp \) is the same set as in 1. above, i.e., \( S^\perp = \text{span}\{(0,1,0),(0,0,1)\} \). We have to include all elements of \( X \) that are orthogonal to the elements of \( S \). That being said, we shall normally be working more along the lines of Example 1, i.e., **subspaces** and their orthogonal complements.

3. Further to the discussion of algebraic vs. orthogonal complementarity, consider the same spaces \( X, Y \) and \( Z \) as defined in Eqs. (20) and (21), but defined over the interval \([-1,1]\). The orthogonal polynomials \( \phi_k(x) \) over \([-1,1]\) that may be constructed from the functions \( x^k, 0 \leq k \leq 10 \) are the so-called **Legendre polynomials**, \( P_n(x) \), listed below:
These polynomials satisfy the following orthogonality property,
\[ \int_{-1}^{1} P_m(x)P_n(x) \, dx = \frac{2}{2n+1} \delta_{mn}. \] (23)

From the above table, we see that the Legendre polynomials \( P_n(x) \), 1 \( \leq \) \( n \) \( \leq \) 5, belong to space \( Y \), whereas the polynomials \( P_n \), 5 \( \leq \) \( n \) \( \leq \) 10, do not belong solely to \( Z \). This suggests that the spaces \( Y \) and \( Z \) are not orthogonal complements. However, the following spaces are orthogonal complements:

\[ Y' = \text{span}\{P_0, P_1, P_2, P_3, P_4, P_5\}, \quad Z' = \text{span}\{P_6, P_7, P_8, P_9, P_{10}\}. \] (24)

(Actually, \( Y' \) is identical to \( Y \) defined earlier.)

There is, however, another decomposition going on in this space, which is made possible by the fact that the interval \([-1, 1]\) is symmetric with respect to the point \( x = 0 \). Note that the polynomials \( P_n(x) \) are either even or odd. This suggests that we should consider the following subsets of \( X \),

\[ Y'' = \{u \in X \mid u \text{ is an even function}\}, \quad Z'' = \{u \in X \mid u \text{ is an odd function}\}. \] (25)

It is a quite simple exercise to show that any function \( f(x) \) defined on an interval \([-a, a]\) may be written as a sum of an even function and an odd function. This implies that any element
u \in X \text{ may be expressed in the form}

\[ u = v + w, \quad v \in Y'', \; w \in Z''. \] (26)

Therefore the spaces \( Y' \) and \( Z' \) are algebraic complements. In terms of the inner product of functions over \([-1,1]\), however, \( Y' \) and \( Z' \) are also orthogonal complements since

\[ \int_{-1}^{1} f(x)g(x) \, dx = 0, \] (27)

if \( f \) and \( g \) have different parity.

The discussion that follows will be centered around Hilbert spaces, i.e., complete inner product spaces. This is because we shall need the closure properties of these spaces, i.e., that they contain the limit points of all sequences. The following result is very important.

The “Projection Theorem” for Hilbert spaces

Let \( H \) be a Hilbert space and \( Y \subset H \) any closed subspace of \( H \). (Note: This means that \( Y \) contains its limit points. In the case of finite-dimensional spaces, e.g., \( \mathbb{R}^n \), a subspace is closed. But a subspace of an infinite-dimensional vector space need not be closed.) Now let \( Z = Y^\perp \). Then for any \( x \in H \), there is a unique decomposition of the form

\[ x = y + z, \quad y \in Y, \; z \in Z = Y^\perp. \] (28)

The point \( y \) is called the (orthogonal) projection of \( x \) on \( Y \).

This is an extremely important result from analysis, and equally important in applications. We’ll examine its implications a little later, in terms of “best approximations” in a Hilbert space.

From Eq. (28, we can define a mapping \( P_Y : H \to Y \), the projection of \( H \) onto \( Y \) so that

\[ P_Y : x \to y. \] (29)

Note that

\[ P_Y : H \to Y, \quad P_Y : Y \to Y, \quad P_Y : Y^\perp \to \{0\}. \] (30)

Furthermore, \( P_Y \) is an idempotent operator, i.e.,

\[ P_Y^2 = P_Y. \] (31)
This follows from the fact that $P_Y(x) = y$ and $P_Y(y) = y$, implying that $P_Y(P_Y(x)) = P_Y(x)$.

Finally, note that
\[ \|x\|^2 = \|y\|^2 + \|z\|^2. \] (32)

This follows from the fact that the norm is defined by means of the inner product:
\[
\|x\|^2 = \langle x, x \rangle = \langle y + z, y + z \rangle \\
= \langle y, y \rangle + \langle z, z \rangle + 2\langle y, z \rangle \\
= \|y\|^2 + \|z\|^2, \] (33)

where the final line results from the fact that $\langle y, z \rangle = \langle z, y \rangle$ since $y \in Y$ and $z \in Z = y^\perp$.

**Example:** Let $H = \mathbb{R}^3$, $Y = \text{span}\{(1,0,0)\}$ and $Y^\perp = \text{span}\{(0,1,0),(0,0,1)\}$. Then $x = (1,2,3)$ admits the unique expansion,
\[
(1,2,3) = (1,0,0) + (0,2,3), \] (34)

where $y = (1,0,0) \in Y$ and $z = (0,2,3) \in Y^\perp$. $y$ is the unique projection of $x$ on $Y$.

**Orthogonal/orthonormal sets of a Hilbert space**

Let $H$ denote a Hilbert space. A set $\{u_1, u_2, \cdots u_n\} \subset H$ is said to form an **orthogonal set** in $H$ if
\[ \langle u_i, u_j \rangle = 0 \quad \text{for} \quad i \neq j. \] (35)

If, in addition,
\[ \langle u_i, u_i \rangle = \|u_i\|^2 = 1, \quad 1 \leq i \leq n, \] (36)

then the $\{u_i\}$ are said to form an **orthonormal set** in $H$.

You will not be surprised by the following result, since you have most probably seen it in earlier courses in linear algebra.

**Theorem:** An orthogonal set $\{u_1, u_2, \cdots\}$ not containing the element $\{0\}$ is linearly independent.

**Proof:** Assume that there are scalars $c_1, c_2, \cdots, c_n$ such that
\[ c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = 0. \] (37)
For each \( k = 1, 2, \cdots n \), form the inner product of both sides of the above equation with \( u_k \), i.e.,

\[
\langle u_k, c_1 u_1 + c_2 u_2 + \cdots + c_n u_n \rangle = \langle u_k, 0 \rangle = 0.
\] (38)

By the orthogonality of the \( u_i \), the LHS of the above equation reduces to \( c_k \| u_k \|_2^2 \), implying that

\[
c_k \| u_k \|_2^2 = 0, \quad k = 1, 2, \cdots, n.
\] (39)

By assumption, however, \( u_k \neq 0 \), implying that \( \| u_k \|_2^2 \neq 0 \). This implies that all scalars \( a_k \) are zero, which means that the set \( \{ u_1, u_2, \cdots, u_n \} \) is linearly independent.

**Note:** As you have also seen in courses in linear algebra, given a linearly independent set \( \{ v_1, v_2, \cdots, v_n \} \), we can always construct an orthonormal set \( \{ e_1, e_2, \cdots, e_n \} \) via the **Gram-Schmidt** orthogonalization procedure. Moreover,

\[
\text{span}\{ v_1, v_2, \cdots, v_n \} = \text{span}\{ e_1, e_2, \cdots, e_n \}.
\] (40)

More on this later.

We have now arrived at the most important result of this section.
Lecture 5

Best approximation in Hilbert spaces

Recall the idea of the best approximation in normed linear spaces, discussed a couple of lectures ago. Let $X$ be an infinite-dimensional normed linear space. Furthermore, let $u_i \in X$, $1 \leq i \leq n$, be a set of $n$ linearly independent elements of $X$ and define the $n$-dimensional subspace,

$$ S_n = \text{span}\{u_1, u_2, \cdots, u_n\}. $$

Then let $x$ be an arbitrary element of $X$. We wish to find the best approximation to $x$ in the subspace $S_n$. It will be given by the element $y_n \in S_n$ that lies closest to $x$, i.e.,

$$ y_n = \arg \min_{v \in S_n} \|x - v\|. $$

(The variables used above may be different from those used in the earlier lecture.)

We are going to use the same idea of best approximation, but in a Hilbert space setting, where we have the additional property that an inner product exists in our space. This, of course, opens the door to the idea of orthogonality, which will play an important role.

The best approximation in Hilbert spaces may be phrased as follows:

**Theorem:** Let $\{e_1, e_2, \cdots, e_n\}$ be an orthonormal set in a Hilbert space $H$. Define $Y = \text{span}\{e_i\}_{i=1}^n$. $Y$ is a subspace of $H$. Then for any $x \in H$, the best approximation of $x$ in $Y$ is given by the unique element

$$ y = P_Y(x) = \sum_{k=1}^n c_k e_k \quad \text{(projection of } x \text{ onto } Y), $$

where

$$ c_k = \langle x, e_k \rangle, \quad k = 1, 2, \cdots, n. $$

The $c_k$ are called the Fourier coefficients of $x$ w.r.t. the set $\{e_k\}$.

Furthermore,

$$ \sum_{k=1}^n |c_k|^2 \leq \|x\|^2. $$

**Proof:** Any element $v \in Y$ may be written in the form

$$ v = \sum_{k=1}^n c_k e_k. $$
The best approximation to $x$ in $Y$ is the point $y \in Y$ that minimizes the distance $\|x - v\|$, $v \in Y$, i.e.,

$$y = \arg \min_{v \in Y} \|x - v\|. \quad (47)$$

In other words, we must find scalars $c_1, c_2, \cdots, c_n$ such that the distance,

$$f(c_1, c_2, \cdots, c_n) = \|x - \sum_{k=1}^{n} c_k e_k\|, \quad (48)$$

is minimized. Here, $f : \mathbb{R}^n$ (or $\mathbb{C}^n$) $\rightarrow \mathbb{R}$. It is easier to consider the non-negative squared distance function,

$$g(c_1, c_2, \cdots, c_n) = \|x - \sum_{k=1}^{n} c_k e_k\|^2 = \langle x - \sum_{k=1}^{n} c_k e_k, x - \sum_{l=1}^{n} c_l e_l \rangle \geq 0. \quad (49)$$

Minimizing $g$ is equivalent to minimizing $f$.

A quick derivation may be done for the real-scalar-valued case, i.e., $c_i \in \mathbb{R}$. Here, an expansion of the inner product on the right yields,

$$g(c_1, c_2, \cdots, c_n) = \|x\|^2 - \sum_{l=1}^{n} c_l \langle x, e_l \rangle - \sum_{k=1}^{n} c_k \langle x, e_k \rangle + \sum_{k=1}^{n} c_k^2. \quad (50)$$

We now impose the necessary stationarity conditions for a minimum,

$$\frac{\partial g}{\partial c_i} = -\langle x, e_i \rangle - \langle e_i, x \rangle + 2c_i = 0, \quad i = 1, 2, \cdots, n. \quad (51)$$

Therefore,

$$c_i = \langle x, e_i \rangle, \quad i = 1, 2, \cdots, n, \quad (52)$$

which identifies a unique point $c \in \mathbb{R}^n$, therefore a unique element $y \in Y$. In order to check that this point corresponds to a minimum, we examine the second partial derivatives,

$$\frac{\partial^2 g}{\partial c_j \partial c_i} = 2\delta_{ij}. \quad (53)$$

In other words, the Hessian matrix is diagonal and positive definite. Therefore the point corresponds to a minimum.

The complex-scalar case, i.e., $c_k \in \mathbb{C}$, is slightly more complicated. A proof (which, of course, would include the real-valued case) is given at the end of this day’s lecture notes.
Finally, substitution of these (real or complex) values of \( c_k \) into the squared distance function in Eq. (49) yields the result

\[
g(c_1, c_2, \cdots c_n) = \|x\|^2 - \sum_{l=1}^{n} |c_l|^2 - \sum_{k=1}^{n} |c_k|^2 + \sum_{k=1}^{n} |c_k|^2
\]

\[
= \|x\|^2 - \sum_{k=1}^{n} |c_k|^2
\]

\[
\geq 0,
\]

which then implies Eq. (45).

Some additional comments regarding the best approximation:

1. The above result implies that the element \( x \in H \) may be expressed uniquely as

\[
x = y + z, \quad y \in Y, \quad z \in Z = Y^\perp.
\]

To see this, define

\[
z = x - y = x - \sum_{k=1}^{n} c_k e_k,
\]

where the \( c_k \) are given by (44). For \( l = 1, 2, \cdots, n \), take the inner product of \( e_l \) with both sides of this equation to give

\[
\langle z, e_l \rangle = \langle x, e_l \rangle - \sum_{k=1}^{n} c_k \langle e_k, e_l \rangle
\]

\[
= \langle x, e_l \rangle - c_l
\]

\[
= 0.
\]

Therefore, \( z \perp e_l, \; l = 1, 2, \cdots, n \), implying that \( z \in Y^\perp \).

2. The term \( z \) in (55) may be viewed as the residual in the approximation \( x \approx y \). The norm of this residual is the magnitude of the error of the approximation \( x \approx y \), i.e.,

\[
\Delta_n = \|z\| = \|x - y\|.
\]

Also recall that \( \Delta_n \) is the distance between \( x \) and the space \( S_n \).

3. As the dimension \( n \) of the orthonormal set \( \{e_1, e_2, \cdots, e_n\} \) is increased, we expect to obtain better approximations to the element \( x \) – unless, of course, \( x \) is an element of one of these finite-dimensional spaces, in which case we arrive at zero approximation error, and no further
improvement is possible. Let us designate \( Y_n = \text{span}\{e_1, e_2, \ldots, e_n\} \), and \( y_n = P_{Y_n}(x) \) the best approximation to \( x \) in \( Y_n \). Then the error of this approximation is given by

\[
\|z_n\| = \|x - y_n\|. \tag{59}
\]

We expect that \( \|z_{n+1}\| \leq \|z_n\| \).

4. Note also that the inequality

\[
\sum_{k=1}^{n} |c_k|^2 \leq \|x\|^2 \tag{60}
\]

holds for all appropriate values of \( n > 0 \). (If \( H \) is finite-dimensional, i.e., \( \text{dim}(H) = N \), then \( n = 1, 2, \ldots, N \).) In other words, the partials sums on the left are bounded from above. This inequality, known as Bessel’s inequality, will have important consequences.

Appendix: Best approximation in the complex-scalar case

We now consider the squared distance function \( g(a) \) in Eq. (49) in the case that \( H \) is a complex Hilbert space, i.e., \( a \in \mathbb{C}^n \).

\[
g(c_1, c_2, \ldots, c_n) = \|x - \sum_{k=1}^{n} c_k e_k\|^2
\]

\[
= \langle x - \sum_{k=1}^{n} c_k e_k, x - \sum_{l=1}^{n} c_l e_l \rangle
\]

\[
= \|x\|^2 - \sum_{k=1}^{n} c_k e_k, x - \sum_{l=1}^{n} c_l e_l \rangle + \sum_{k=1}^{n} \sum_{l=1}^{n} \langle c_k e_k, c_l e_l \rangle
\]

\[
= \|x\|^2 - \sum_{k=1}^{n} \left[ \overline{a}_k \langle x, e_k \rangle + c_k \langle x, e_k \rangle + |c_k|^2 \right]
\]

\[
= \|x\|^2 + \sum_{k=1}^{n} \left[ |\langle x, e_k \rangle|^2 - \overline{a}_k \langle x, e_k \rangle - c_k \langle x, e_k \rangle + |c_k|^2 \right] - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2
\]

\[
= \|x\|^2 + \sum_{k=1}^{n} |\langle x, e_k \rangle - c_k|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2. \tag{61}
\]

The first and last terms are fixed. The middle term is a sum of nonnegative numbers. The minimum value is achieved when all of these terms are zero. Consequently, \( f(c_1, \ldots, c_n) \) is a minimum if and only if \( c_k = \langle x, e_k \rangle \) for \( k = 1, 2, \ldots, n \). As in the real case, we have

\[
\|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \geq 0. \tag{62}
\]
Some examples

We now consider some examples to illustrate the property that the best approximation of an element \( x \in H \) in a subspace \( Y \subset H \) is the projection \( P_Y(x) \).

1. The finite dimensional case \( H = \mathbb{R}^3 \), as an easy “starter”. Let \( x = (a, b, c) = a \mathbf{e}_1 + b \mathbf{e}_2 + c \mathbf{e}_3 \).

Geometrically, the \( \mathbf{e}_i \) may be visualized as the \( \mathbf{i} \), \( \mathbf{j} \) and \( \mathbf{k} \) unit vectors which form an orthonormal set. Now let \( Y = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\} \). Then \( y = P_Y(x) = (a, b, 0) \), which lies in the \( \mathbf{e}_1 \)-\( \mathbf{e}_2 \) plane. Moreover, the distance between \( x \) and \( y \) is

\[
||x - y|| = [(a - a)^2 + (b - b)^2 + (c - 0)^2]^{1/2} = |c|,
\]

the distance from \( y \) to the \( \mathbf{e}_1 \)-\( \mathbf{e}_2 \) plane.

2. Now let \( H = L^2[-\pi, \pi] \), the space of square-integrable functions on \([−\pi, \pi]\) and consider the following set of functions,

\[
e_1(x) = \frac{1}{\sqrt{2\pi}}, \quad e_2(x) = \frac{1}{\sqrt{\pi}} \cos x, \quad e_3(x) = \frac{1}{\sqrt{\pi}} \sin x.
\]

These three functions form an orthonormal set in \( H \), i.e., \( \langle e_i, e_j \rangle = \delta_{ij} \). Let \( Y_3 = \text{span}\{e_1, e_2, e_3\} \).

Now consider the function \( f(x) = x^2 \) in \( L^2[-\pi, \pi] \). The best approximation to \( f \) in \( Y_3 \) will be given by the function

\[
f_3 = P_{Y_3} f = \langle f, e_1 \rangle e_1 + \langle f, e_2 \rangle e_2 + \langle f, e_3 \rangle e_3.
\]

We now compute the Fourier coefficients:

\[
\langle f, e_1 \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x^2 \, dx = \cdots = \frac{\sqrt{2}}{3} \pi^{5/2},
\]

\[
\langle f, e_2 \rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x^2 \cos x \, dx = \cdots = -4\sqrt{\pi}
\]

\[
\langle f, e_3 \rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x^2 \sin x \, dx = 0.
\]

The final result is

\[
f_3(x) = \frac{\sqrt{2}}{3} \pi^{5/2} \left( \frac{1}{\sqrt{2\pi}} \right) - 4\sqrt{\pi} \left( \frac{1}{\sqrt{\pi}} \cos x \right)
\]

\[
= \frac{\pi^2}{3} - 4 \cos x.
\]
Note: This result is identical to the result obtained by the traditional Fourier series method, where you simply worked with the \( \cos x \) and \( \sin x \) functions and computed the expansion coefficients using the formulas from AMATH 231, cf. Lecture 2. Computationally, more work is involved in producing the above result because you have to work with factors and powers of \( \sqrt{\pi} \) that may eventually disappear. The advantage of the above approach is to illustrate the “best approximation” idea in terms of projections onto spans of orthonormal sets.

Finally, we compute the error of the above approximation to be (via MAPLE)

\[
\| f - f_3 \|_2 = \left[ \int_{-\pi}^{\pi} \left( x^2 - \frac{\pi^3}{3} + 4 \cos x \right)^2 \, dx \right]^{1/2} \approx 2.034. \tag{68}
\]

The approximation is sketched in the top figure on the next page.

3. Same space and function \( f(x) = x^2 \) as above, but we add two more elements to our orthonormal basis set,

\[
e_4(x) = \frac{1}{\sqrt{\pi}} \cos 2x, \quad e_5(x) = \frac{1}{\sqrt{\pi}} \sin 2x. \tag{69}
\]

Now define \( Y_5 = \text{span}\{e_1, \ldots, e_5\} \), We shall denote the approximation to \( f \) in this space as \( f_5(x) = P_{Y_5} f \),

\[
f_5 = \sum_{k=1}^{5} \langle f, e_k \rangle e_k = f_3 + \langle f, e_4 \rangle e_4 + \langle f, e_5 \rangle e_5. \tag{70}
\]

In other words, as you already know, the first three coefficients of the expansion do not have to be recomputed. The final two Fourier coefficients are now computed,

\[
\langle f, e_4 \rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x^2 \sin 2x \, dx = 0
\]

\[
\langle f, e_5 \rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x^2 \cos 2x \, dx = \pi. \tag{71}
\]

The final result is

\[
f_5(x) = \frac{\pi^2}{3} - 4 \cos x + \sqrt{\pi} \cos 2x. \tag{72}
\]

This approximation is sketched in the bottom figure on the next page. Finally, the error of this approximation is computed to be (via MAPLE)

\[
\| f - f_5 \|_2 \approx 1.694, \tag{73}
\]

which is lower than the approximation error yielded by \( f_3 \), as expected.
Approximation $f_3(x) = (P_{Y_3}f)(x)$ to $f(x) = x^2$. Error $\|f - f_3\|_2 \approx 2.034$.

Approximation $f_5(x) = (P_{Y_5}f)(x)$ to $f(x) = x^2$. Error $\|f - f_5\|_2 \approx 1.694$. 
4. Now consider the space $H = L^2[-1, 1]$ and the following set of functions

$$e_1(x) = \frac{1}{\sqrt{2}}, \quad e_2(x) = \sqrt{\frac{3}{2}} x, \quad e_3(x) = \sqrt{\frac{5}{2} \cdot \frac{1}{2}} (3x^2 - 1).$$  \tag{74}

These three functions form an orthonormal set on $[-1, 1]$. Moreover, span$\{e_1, e_2, e_3\} = $ span$\{1, x, x^2\}$. These functions result from the application of the Gram-Schmidt orthogonalization procedure to the linearly independent set $\{1, x, x^2\}$.

5. We now consider a slightly more intriguing example that will provide a preview to our study of wavelet functions later in this course. Consider the function space $L^2[0, 1]$ and the two elements $e_1$ and $e_2$ given by

$$e_1(x) = 1, \quad e_2(x) = \begin{cases} 
1, & 0 \leq x \leq 1/2, \\
-1, & 1/2 < x \leq 1.
\end{cases}$$  \tag{75}

They are sketched in the figure below.

It is not too hard to see that these two functions form an orthonormal set in $L^2[0, 1]$, i.e.,

$$\langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = 1, \quad \langle e_1, e_2 \rangle = 0.$$  \tag{76}

Now let $f(x) = x^2$ as before. Let us first consider the subspace $Y_1 = \text{span}\{e_1\}$. It is the one-dimensional subspace of functions in $L^2[0, 1]$ that are constant on the interval. The approximation to $f$ in this space is given by

$$f_1 = (P_{Y_1}) f = \langle f, e_1 \rangle e_1 = \langle f, e_1 \rangle,$$  \tag{77}

since $e_1 = 1$. The Fourier coefficient is given by

$$\langle f, e_1 \rangle = \int_0^1 x^2 \, dx = \frac{1}{3}.$$  \tag{78}

Therefore, the function $f_1(x) = \frac{1}{3}$, sketched in the left subfigure below, is the best constant-function approximation to $f(x) = x^2$ on the interval. It is the mean value of $f$ on $[0, 1]$. 48
Now consider the space $Y_2 = \text{span}\{e_1, e_2\}$. The best approximation to $f$ in this space will be given by

$$f_2 = \langle f, e_1 \rangle e_1 + \langle f, e_2 \rangle e_2. \quad (79)$$

The first term, of course, has already been computed. The second Fourier coefficient is given by

$$\langle f, e_2 \rangle = \int_0^1 x^2 e_2(x) \, dx = \int_0^{1/2} x^2 \, dx - \int_{1/2}^1 x^2 \, dx = \frac{1}{3} x^3 \bigg|_0^{1/2} - \frac{1}{3} x^3 \bigg|_{1/2} = -\frac{1}{4}. \quad (80)$$

Therefore

$$f_2 = \frac{1}{3} e_1 - \frac{1}{4} e_2. \quad (81)$$

In order to get the graph of $f_2$ from that of $f_1$, we simply subtract $1/4$ from the value of $1/3$ over the interval $[0, 1/2]$ and add $1/4$ to the value of $1/3$ over the interval $(1/2, 1]$. The result is

$$f_3(x) = \begin{cases} 
1/12, & 0 \leq x \leq 1/2, \\
7/12, & 1/2 < x \leq 1.
\end{cases} \quad (82)$$

The graph of $f_3(x)$ is sketched in the right subfigure below. The values $1/12$ and $7/12$ correspond to the mean values of $f(x) = x^2$ over the intervals $[0, 1/2]$ and $(1/2, 1]$, respectively. (These should, of course, agree with your calculations in Problem Set No. 1.)

The space $Y_2$ is the vector space of functions in $L^2[0, 1]$ that are piecewise-constant over the half-intervals $[0, 1/2]$ and $(1/2, 1]$. The function $f_2$ is the best approximation to $x^2$ from this space.
A natural question to ask is, “What would be the next functions in this set of piecewise constant orthonormal functions?” Two possible candidates are the functions sketched below.

These, in fact, are called “Walsh functions” and have been used in signal image processing. However, another set of functions which can be employed, and which will be quite relevant later in the course, include the following:

These are the next two “Haar wavelet functions”. We claim that the space $Y_4 = \text{span}\{e_1, e_2, e_3, e_4\}$ is the set of all functions in $L^2[0, 1]$ that are piece-wise constant on the half-intervals $[0, 1/2]$ and $(1/2, 1]$.

**A note on the Gram-Schmidt orthogonalization procedure**

As you may recall from earlier courses in linear algebra, the Gram-Schmidt procedure allows the construction of an orthonormal set of elements $\{e_k\}_1^n$ from a linearly independent set $\{v_k\}_1^n$, with $\text{span}\{e_k\}_1^n = \text{span}\{v_k\}_1^n$. Here we simply recall the procedure.

Start with an element, say $v_1$, and define

$$e_1 = \frac{v_1}{\|v_1\|}. \quad (83)$$
Now take element \( v_2 \) and remove the component of \( e_1 \) from \( v_2 \) by defining

\[
z_2 = v_2 - \langle v_2, e_1 \rangle e_1.
\]  
(84)

We check that \( e_1 \perp z_2 \):

\[
\langle e_1, z_2 \rangle = \langle e_1, v_2 \rangle - \langle e_1, v_2 \rangle \langle e_1, e_1 \rangle = 0.
\]  
(85)

Now define

\[
e_2 = \frac{z_2}{\|z_2\|}.
\]  
(86)

We continue the procedure, taking \( v_3 \) and eliminating the components of \( e_1 \) and \( e_2 \) from it. Define,

\[
z_3 = v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2.
\]  
(87)

It is straightforward to show that \( z_3 \perp e_1 \) and \( e_2 \). Then define

\[
e_3 = \frac{z_3}{\|z_3\|}.
\]  
(88)

In general, from a knowledge of \( \{e_1, \cdots, e_{k-1}\} \), we can produce the next element \( e_k \) as follows:

\[
e_k = \frac{z_k}{\|z_k\|}, \text{ where } z_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i.
\]  
(89)

Of course, if the inner product space in which we are working is finite dimensional, then the procedure terminates at \( k = n = \dim(H) \). But when \( H \) is infinite-dimensional, we may, at least in principle, be able to continue to process indefinitely, producing a countably infinite orthonormal set of elements \( \{e_k\}_{1}^{\infty} \). The next question is, “Is such an orthonormal set useful?” The answer is, “Yes.”
Lecture 6

Inner product spaces (cont’d)

Complete orthonormal basis sets in an infinite-dimensional Hilbert space

Let $H$ be an infinite-dimensional Hilbert space. Let us also suppose that we have an infinite sequence of orthonormal elements $\{e_k\} \subset H$, with $\langle e_i, e_j \rangle = \delta_{ij}$. We consider the finite orthonormal sets $E_n = \{e_1, e_2, \ldots, e_n\}, n = 1, 2, \ldots$, and define

$$V_n = \text{span}\{e_1, e_2, \ldots, e_n\}, \quad n = 1, 2, \ldots. \quad (90)$$

Clearly, each $V_n$ is an $n$-dimensional subspace of $H$.

Recall that for an $x \in H$, the best approximation to $x$ in $V_n$ is given by

$$y_n = P_{V_n}(x) = \sum_{k=1}^{n} \langle x, e_k \rangle e_k, \quad (91)$$

with

$$\|y_n\|^2 = \sum_{k=1}^{n} |\langle x, e_k \rangle|^2. \quad (92)$$

Let us denote the error associated with the approximation $x \approx y_n$ as

$$\Delta_n = \|x - y_n\|. \quad (93)$$

Note that this error is the distance between $x$ and $y_n$ as defined by the norm on $H$ which, in turn, is defined by the inner product $\langle \cdot, \cdot \rangle$ on $H$.

Now consider $V_{n+1} = \text{span}\{e_1, e_2, \ldots, e_n, e_{n+1}\}$, which we may write as

$$V_{n+1} = V_n \oplus \text{span}\{e_{n+1}\}. \quad (94)$$

It follows that

$$V_n \subset V_{n+1}. \quad (95)$$

This, in turn implies that

$$\Delta_{n+1} \leq \Delta_n. \quad (96)$$

We can achieve the same error $\Delta_n$ in $V_{n+1}$ by imposing the condition that the coefficient $c_{n+1}$ of $e_{n+1}$ is zero in the approximation. By allowing $c_{n+1}$ to vary, it might be possible to obtain a better
approximation. In other words, since we are minimizing over a larger set, we can’t do any worse than we did before.

If our Hilbert space $H$ were finite dimensional, i.e., $\dim(H) = N > 0$, then $\Delta_N = 0$ for all $x \in H$. But in the case that $H$ is infinite-dimensional, we would like that

$$\Delta_n \to 0 \text{ as } n \to \infty, \text{ for all } x \in H.$$  (97)

In other words all approximation errors go to zero in the limit. (Of course, in the particular case that $x \in V_N$, then $\Delta_N = 0$. But we want to be able to say something about all $x \in H$.) Then we shall be able to write the infinite-sum result,

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k.$$  (98)

The property (97) will hold provided that the orthonormal set $\{e_k\}^\infty_1$ is a complete or maximal orthonormal set in $H$.

**Definition:** An orthonormal set $\{e_k\}^\infty_1$ is said to be complete or maximal if the following is true:

If $\langle x, e_k \rangle = 0$ for all $k \geq 1$ then $x = 0$.  (99)

The idea is that the $\{e_k\}$ elements “detect everything” in the Hilbert space $H$. And if none of them detect anything in an element $x \in H$, then $x$ must be the zero element.

Now, how do we know if a complete orthonormal set can exist in a given Hilbert space? The answer is that if the Hilbert space is separable, then such a complete, countably-infinite set exists. (A separable space contains a dense countable subset.) OK, so this doesn’t help yet, because we now have to know whether our Hilbert space of interest is separable. Let it suffice here to state that most of the Hilbert spaces that we use in applications are separable. (See Note below.) Therefore, complete orthonormal basis sets can exist. And the final “icing on the cake” is the fact that, for separable Hilbert spaces, the Gram-Schmidt orthogonalization procedure can produce such a complete orthonormal basis.

**Note to above:** In the case of $L^2[a, b]$, we have the following results from advanced analysis: (i) The space of all polynomials $\mathcal{P}[a, b]$ defined on $[a, b]$ is dense in $L^2[a, b]$. That means that given any function $u \in L^2[a, b]$ and an $\epsilon > 0$, we can find an element $p \in \mathcal{P}[a, b]$ such that $\|u - p\|_2 < \epsilon$. (ii) The set of
polynomials with rational coefficients, call it $P_R[a, b]$, a subset of $P[a, b]$ is dense in $P[a, b]$. (You may know that the set of rational numbers is dense in $\mathbb{R}$.) And finally, the set $P[a, b]$ is countable. (The set of rational numbers on $\mathbb{R}$ is countable.) (iii) Therefore the set $P_R[a, b]$ is a dense and countable subset of $L^2[a, b]$.

**Complete orthonormal basis sets – “Generalized Fourier series”**

We now conclude our discussion of complete orthonormal basis sets in a Hilbert space.

In what follows, we let $H$ denote an infinite-dimensional Hilbert space (for example, the space of square-integrable functions $L^2[-\pi, \pi]$). It may help to recall the following definition.

**Definition:** An orthonormal set $\{e_k\}_{1}^{\infty}$ is said to be complete or maximal if the following is true:

$$\langle x, e_k \rangle = 0 \text{ for all } k \geq 1 \text{ then } x = 0.$$  \hspace{1cm} (100)

Here is the main result:

**Theorem:** Let $\{e_k\}_{1}^{\infty}$ denote an orthonormal set on a separable Hilbert space. Then the following statements are equivalent:

1. The set $\{e_k\}_{1}^{\infty}$ is complete (or maximal). (In other words, it serves as a complete basis for $H$.)

2. For any $x \in H$,

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k.$$  \hspace{1cm} (101)

(In other words, $x$ has a unique representation in the basis $\{e_k\}$.)

3. For any $x \in H$,

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2.$$  \hspace{1cm} (102)

This is called Parseval’s equation.

**Notes:**

1. The expansion in Eq. (101) is also called a “Generalized Fourier Series”. Note that the basis elements $e_k$ do not have to be sine or cosine functions – they can be polynomials in $x$: the term “Generalized Fourier Series” may still be used.
2. The coefficients $c_k = \langle x, e_k \rangle$ in Eq. (101) are often called “Fourier coefficients,” once again, even if the $e_k$ are not sine or cosine functions.

3. Perhaps the most important note to be made concerns Eq. (102): The sequence $c = (c_1, c_2, \cdots)$ is seen to be square-summable. In other words, $a \in l^2$: $a$ is an element of the sequence space $l^2$. In fact, we may rewrite Eq. (102) as

$$\|x\|_{L^2} = \|c\|_{l^2}.$$  \hspace{1cm} (103)

An important consequence of the above theorem:

As before, let $V_n = \text{span}\{e_1, e_2, \cdots, e_n\}$. For a given $x \in H$, let $y_n \in V_n$ be the best approximation of $x$ in $V_n$ so that

$$y_n = \sum_{k=1}^{n} c_k e_k, \quad c_k = \langle x, e_k \rangle.$$  \hspace{1cm} (104)

Then the magnitude of the error of approximation $x \approx y_n$ is given by

$$\Delta_n = \|x - y_n\| = \left\| \sum_{k=1}^{\infty} c_k e_k - \sum_{k=1}^{n} c_k e_k \right\| = \left\| \sum_{k=n+1}^{\infty} c_k e_k \right\| = \left[ \sum_{k=n+1}^{\infty} c_k^2 \right]^{1/2}. \hspace{1cm} (105)$$

In other words, the magnitude of the error is the magnitude (in $l^2$ norm) of the “tail” of the sequence of Fourier coefficients $c$, i.e., the sequence of Fourier coefficients $\{c_{n+1}, c_{n+2}, \cdots\}$ that has been “thrown away” by the approximation $x \approx y_n$. This truncation of the infinite sequence of Fourier coefficients $c$ occurs because the basis functions $e_{n+1}, e_{n+2}, \cdots$ do not belong to $V_n$ – we are considering only those coefficients $c_k$ corresponding to basis elements $e_k \in V_n$.

We shall be returning to this idea frequently in the course.
Some alternate versions of Fourier sine/cosine series

1. We have already seen and used one important orthonormal basis set: The (normalized) cosine/sine basis functions for Fourier series on \([-\pi, \pi]\):

\[ \{e_k\} = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin x, \frac{1}{\sqrt{\pi}} \cos 2x, \cdots \right\} \]  \hspace{1cm} (106)

These functions form a complete basis in the space of real-valued square-integrable functions \(L^2[-\pi, \pi]\).

This is a natural, but particular case, of the more general class of orthonormal sine/cosine functions on the interval \([-a, a]\), where \(a > 0\):

\[ \{e_k\} = \left\{ \frac{1}{\sqrt{2a}}, \frac{1}{\sqrt{a}} \cos \left( \frac{\pi x}{a} \right), \frac{1}{\sqrt{a}} \sin \left( \frac{\pi x}{a} \right), \frac{1}{\sqrt{a}} \cos \left( \frac{2\pi x}{a} \right), \cdots \right\} \]  \hspace{1cm} (107)

These functions form a complete basis in the space of real-valued square-integrable functions \(L^2[-a, a]\). When \(a = \pi\), we have the usual Fourier series functions \(\cos kx\) and \(\sin kx\). We shall return to this set in the next lecture.

2. Sometimes, it is convenient to employ the following set of complex-valued square-integrable functions on \([-\pi, \pi]\):

\[ e_k = \frac{1}{\sqrt{\pi}} \exp(ikx), \quad k = \cdots, -2, -1, 0, 1, 2 \cdots . \]  \hspace{1cm} (108)

Note that the index \(k\) is infinite in both directions. These functions form a complete set in the complex-valued space \(L^2[-\pi, \pi]\). The orthonormality of this set with respect to the complex-valued inner product on \([-\pi, \pi]\) is left as an exercise. Because of Euler’s formula,

\[ e^{ikx} = \cos kx + i \sin kx, \]  \hspace{1cm} (109)

expansions in this basis are related to Fourier series expansions. We shall return to this set in the near future.

By means of scaling, the following set forms a basis for the complex-valued space \(L^2[-a, a]\):

\[ e_k = \frac{1}{\sqrt{a}} \exp \left( \frac{ik\pi x}{a} \right), \quad k = \cdots, -2, -1, 0, 1, 2 \cdots . \]  \hspace{1cm} (110)

3. The following functions

\[ \{e_k\} = \left\{ 1, \sqrt{2} \cos \pi x, \sqrt{2} \sin \pi x, \sqrt{2} \cos 2\pi x, \sqrt{2} \sin 2\pi x, \cdots \right\} . \]  \hspace{1cm} (111)

form an orthonormal set on the space \(L^2[0, 1]\).
Convergence of Fourier series expansions

Here we briefly discuss some convergence properties of Fourier series expansions, namely,

1. pointwise convergence – the convergence of a series at a point \( x \),

2. uniform convergence – the convergence of a series on an interval \( [a, b] \) in \( \| \|_\infty \) norm/metric,

3. convergence in mean – the convergence of a series on an interval \( [a, b] \) in \( \| \|_2 \) norm/metric.

You may have seen some, perhaps all, of these ideas in AMATH 231. We shall cover them briefly, and without proof. What will be of greater concern to us is the rate of convergence of a series and its importance in signal/image processing.

Recall that the Fourier series expansion of a function \( f(x) \) over the interval \( [-\pi, \pi] \) had the following form,

\[
f(x) = a_0 + \sum_{k=1}^{\infty} [a_k \cos kx + b_k \sin kx], \quad x \in [-\pi, \pi].
\] (112)

The right-hand-side of Eq. (112) is clearly a 2\( \pi \)-periodic function of \( x \). As such, one expects that it will represent a 2\( \pi \)-periodic function. Indeed, as you probably saw in AMATH 231, this is the case. If we consider values of \( x \) outside the interval \( (-\pi, \pi) \), then the series will represent the so-called “2\( \pi \)-extension” of \( f(x) \) – one essentially takes the graph of \( f(x) \) on \( (-\pi, \pi) \) and copies it on each interval \( (-\pi + 2k\pi, \pi + 2k\pi) \), \( k = -1, -2, \cdots \) and \( k = 1, 2, \cdots \). There are some potential complications, however, at the connection points \( (2k+1)\pi, k \in \mathbb{R} \). It is for this reason that we used the open interval \( (-\pi, \pi) \) above.

**Case 1:** \( f \) is 2\( \pi \)-periodic. In this case, there are no problems with translating the graph of \( f(x) \) on \( [-\pi, \pi] \), since \( f(-\pi) = f(\pi) \). Two contiguous graphs will intersect at the connection points. Without loss of generality, let us simply assume that \( f(x) \) is continuous on \( [-\pi, \pi] \). Then the graph of its 2\( \pi \)-extension is continuous at all \( x \in \mathbb{R} \). A sketch is given below.

**Case 2:** \( f \) is not 2\( \pi \)-periodic. In particular \( f(-\pi) \neq f(\pi) \). Then there is no way that two contiguous graphs of \( f(x) \) will intersect at the connection points – there will be discontinuities, as sketched below.

The existence of such discontinuities in the 2\( \pi \)-extension of \( f \) will have consequences regarding the convergence of Fourier series to the function \( f(x) \), even on \( (-\pi, \pi) \). Of course, there may be other discontinuities of \( f \) inside the interval \( (-\pi, \pi) \), which will also affect the convergence.
Case 1: 2π-extension of a 2π-periodic function \( f(x) \).

We now state some convergence results, starting with the “weakest result,” i.e., the result that has minimal assumptions on \( f(x) \).

**Convergence Result No. 1:** \( f \) is square-integrable on \([−π, π]\). Mathematically, we require that \( f \in L^2[−π, π]\), that is,

\[
\int_{−π}^{π} |f(x)|^2 \, dx < \infty. \tag{113}
\]

As we discussed in a previous lecture, the function \( f(x) \) doesn’t have to be continuous – it can be piecewise continuous, or even worse! For example, it doesn’t even have to be bounded. In the applications examined in this course, however, we shall be dealing with bounded functions. In engineering parlance, if \( f \) satisfies the condition in Eq. (113) then it is said to have “finite energy.”

In this case, the convergence result is as follows: The Fourier series in (112) converges in mean to \( f \). In other words, the Fourier series converges to \( f \) in \( L^2 \) norm/metric. By this we mean that the partial sums \( S_n \) of the Fourier series converge to \( f \) as follows,

\[
\|f - S_n\|_2 \to 0 \quad \text{as} \quad n \to \infty. \tag{114}
\]

Case 2: 2π-extension of a function \( f(x) \) that is not 2π-periodic.
Note that this property does not imply that the partial series $S_n$ converge pointwise, i.e., that $|f(x) - S_n(x)| \to \infty$ as $n \to \infty$ for any $x$.

As for a proof of this convergence result: It follows from the Theorem stated at the beginning of this lecture. The sine/cosine basis used in Fourier series is complete in $L^2[-\pi, \pi]$.

On the other side of the spectrum, we have the “strongest result,” i.e., the result that has quite stringent demands on the behaviour of $f$.

**Convergence Result No. 2:** $f$ is $2\pi$-periodic, continuous and piecewise $C^1$ on $[-\pi, \pi]$. In this case, the Fourier series in (112) converges uniformly to $f$ on $[-\pi, \pi]$. This means that the Fourier series converges to $f$ in the $\|\cdot\|_\infty$ norm/metric: From previous discussions, this implies that the partial sums $S_n$ converge to $f$ as follows,

$$\|f - S_n\|_\infty \to 0 \quad \text{as} \quad n \to \infty. \quad (115)$$

This is a very strong result – it implies that the partial sums $S_n$ converge pointwise to $f$, i.e.,

$$|f(x) - S_n(x)| \to 0 \quad \text{as} \quad n \to \infty. \quad (116)$$

But the result is actually stronger than this since the pointwise convergence is uniform over the interval $[-\pi, \pi]$, in an “$\varepsilon$-ribbonlike” fashion. This comes from the definition of the $\|\cdot\|_\infty$ norm.

A proof of this result can be found in the book by Boggess and Narcowich – see Theorem 1.30 and its proof, pp. 72-75.

**Example:** Consider the function

$$f(x) = |x| = \begin{cases} -x, & -\pi < x \leq 0, \\ x, & 0 < x \leq \pi, \end{cases} \quad (117)$$

which is continuous on $[-\pi, \pi]$. Moreover, its $2\pi$-extension is also continuous on $\mathbb{R}$ since $f(-\pi) = f(\pi) = \pi$. Because the function $f(x)$ is even, the expansion is only in terms of the cosine functions.

The series has the form (Exercise)

$$f(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos kx, \quad a_0 = \frac{\pi}{2}, \quad a_k = \begin{cases} -\frac{4}{k^2}, & k \ \text{odd}, \\ 0, & k \ \text{even}, \end{cases} \quad (118)$$
In the figure below is presented a plot of the partial sum $S_9(x)$ which is comprised of only six nonzero coefficients, $a_0, a_1, a_3, a_5, a_7, a_9$. Despite the fact that we use only six terms of the Fourier series, an excellent approximation to $f(x)$ is achieved over the interval $[-\pi, \pi]$. The use of 11 nonzero coefficients, i.e., the partial sum $S_{19}(x)$ produces an approximation that is virtually indistinguishable from the plot of the function $f(x)$ in the figure!

Partial sum $S_9(x)$ of the Fourier cosine series expansion (118) of the $2\pi$-periodic piecewise constant function,

$$f(x) = |x| = \begin{cases} -x, & -\pi < x \leq 0, \\ x, & 0 < x \leq \pi, \end{cases}$$

(119)

The function $f(x)$ is also plotted.

Convergence Result No. 2 is applicable in this case, so we may conclude that the Fourier series converges uniformly to $f(x)$ over the entire interval $[-\pi, \pi]$. That being said, we notice that the degree of accuracy achieved at the points $x = 0, x = \pm \pi$ is not the same as at other points, in particular, $x = \pm \pi/2$. Even though uniform convergence is guaranteed, the rate of convergence is seen to be a little slower at these “kinks”. These points actually represent singularities of the function – not points of discontinuity of the function $f(x)$ but of its derivative $f'(x)$. Even such singularities can affect the rate of convergence of a Fourier series expansion. We’ll say more about this later.
Uniform convergence implies convergence in mean

We expect that if the stronger result, No. 2, applies to a function \( f \), then the weaker result will also apply to it, i.e.,

uniform convergence on \([-\pi, \pi]\) implies \( L^2 \) convergence on \([-\pi, \pi]\).

A quick way to see this is that if \( f \in C[a,b] \), it is bounded on \([a,b]\), implying that it must be in \( L^2[a,b] \). But let’s go through the mathematical details, since they are revealing.

Suppose that \( f \in C[a,b] \). Its \( L^2 \) norm is given by

\[
\| f \|_2 = \left( \int_a^b |f(x)|^2 \, dx \right)^{1/2}.
\] (120)

Since \( f \in C[a,b] \) it is bounded on \([a,b]\). Let \( M \) denote the value of the infinity norm of \( f \), i.e.,

\[
M = \max_{a \leq x \leq b} |f(x)| = \| f \|_\infty.
\] (121)

Now return to Eq. (120) and note that, from the basic properties of integrals,

\[
\int_a^b |f(x)|^2 \, dx \leq \int_a^b M^2 \, dx = M^2 (b-a).
\] (122)

Substituting this result into (120), we have

\[
\| f \|_2 \leq M \sqrt{b-a} = \sqrt{b-a} \| f \|_\infty.
\] (123)

Now replace \( f \) with \( f - S_n \):

\[
\| f - S_n \|_2 \leq \sqrt{b-a} \| f - S_n \|_\infty.
\] (124)

Uniform convergence implies that the RHS goes to zero as \( n \to \infty \). This, in turn, implies that the LHS goes to zero as \( n \to \infty \), which implies convergence in \( L^2 \), proving the desired result.

Convergence Results 1 and 2 appear to represent opposite sides of the spectrum, in terms of the behaviour of \( f \). Result 1 assumes very little from \( f \), whereas Result 2 assumes a good deal, i.e., continuity. The following result, where \( f \) is assumed to be piecewise continuous, is a kind of intermediate result which is quite applicable in signal and image processing. Recall that \( f \) is said to be piecewise continuous on an interval \( I \) if it is continuous at all \( x \in I \) with the exception of a finite number of points in \( I \). In this way, it can have “jumps”.

61
Convergence Result No. 3: $f$ is piecewise $C^1$ on $[-\pi, \pi]$. In this case:

1. The Fourier series converges uniformly to $f$ on any closed interval $[a, b]$ that does not contain a point of discontinuity of $f$.

2. If $p$ denotes a point of discontinuity of $f$, then the Fourier series converges to the value

$$\frac{f(p + 0) + f(p - 0)}{2},$$

where

$$f(p + 0) = \lim_{h \to 0^+} f(p + h), \quad f(p - 0) = \lim_{h \to 0^+} f(p - h).$$

Notes:

1. The “piecewise $C^1$” requirement, as opposed to “piecewise $C$” guarantees that the slopes $f'(x)$ of tangents to the curve remain finite as $x$ approaches points of discontinuity both from the left and from the right. A proof of this convergence result may be found in the book by Boggess and Narcowich, cf. Theorem 1.22, p. 63 and Theorem 1.28, p. 70.

2. Statement 1. above is stronger than the corresponding result presented in AMATH 231. In Theorem 5.5 (ii), Page 151, AMATH 231 Course Notes, it is stated that if the function $f$ is piecewise $C^1$, then its Fourier series converges pointwise at all $x$.

Example: Consider the function defined by

$$f(x) = \begin{cases} 
-1, & -\pi < x \leq 0, \\
1, & 0 < x \leq \pi.
\end{cases}$$

Because the function $f(x)$ is odd, the expansion is only in terms of the sine functions (as you found in Problem Set No. 1). The series has the form

$$f(x) = \sum_{k=1}^{\infty} b_k \sin kx,$$

$$b_k = \begin{cases} 
\frac{4}{(k\pi)}, & k \text{ odd}, \\
0, & k \text{ even}.
\end{cases}$$

Clearly, $f(x)$ is discontinuous at $x = 0$ because of the jump there. But its $2\pi$-extension is also discontinuous at $x = \pm\pi$. In the figure below is presented a plot of the partial sum $S_{50}(x)$ of the Fourier series expansion to this function.
Partial sum $S_{50}(x)$ of Fourier sine series expansion (128) of the $2\pi$-periodic piecewise constant function,

$$f(x) = \begin{cases} -1, & -\pi < x \leq 0, \\ 1, & 0 < x \leq \pi. \end{cases}$$  \hspace{1cm} (129)$$

The function $f(x)$ is also plotted.

Clearly, $f(x)$ is continuous at all $x \in [-\pi, \pi]$ except at $x = 0$ and $x = \pm \pi$. In the vicinity of these points, the convergence of the Fourier series appears to be slower – one would need a good number of additional terms in the expansion in order to approximate $f(x)$ near these points to the accuracy demonstrated elsewhere, say near $x = \pm \pi/2$. According to the first point of Convergence Result No. 3, the Fourier series converges uniformly on any closed interval $[a, b]$ that does not contain the points of discontinuity $x = 0, x = -\pi, x = \pi$. Even though the convergence on such a closed interval is uniform, it may not necessarily be very rapid. Compare this result to that obtained by only six terms of the Fourier series to the continuous function $f(x) = |x|$ shown in the previous figure. We’ll return to this point in the next lecture.

Intuitively, one may imagine that it takes a great deal of effort for the series to be approximating the value $f(x) = -1$ for negative values of $x$ near 0, and then having to jump up to approximate the value $f(x) = 1$ for positive values of $x$ near 0. As such, more terms of the expansion are required because of the dramatic jump in the function. We shall return to this point in the next lecture as well.

At each of the three discontinuities in the plot, we see that the second point of Convergence Result No. 3, regarding the behaviour of the Fourier series at a discontinuity, is obeyed. For example, at $x = 0$, the series converges to zero, because all terms are zero: $\sin(0) = 0$ for all $k$. And zero is precisely the average value of the left and right limits $f(1 - 0) = -1$ and $f(1 + 0) = 1$. The same holds true at $x = \pi$ and $x = -\pi$. 

63
The visible oscillatory behaviour of the partial sum function $S_{50}(x)$ in the plot is called “Gibbs ringing” or the “Gibbs artifact.” For lower partial sums, i.e., $S_n(x)$ for $n < 50$, the oscillatory nature is even more pronounced. Such “ringing” is a fact-of-life in image processing, since images generally contain a good number of discontinuities, namely edges. Since image compression methods such as JPEG rely on the truncation of Fourier series, they are generally plagued by ringing artifacts near edges. This is yet another point that will be addressed later in this course.