A Brief Introduction to Hilbert Space

Yang Chen*, Shaoshu Li†

December, 2016

1 Introduction

We have been familiar with the concept of linear vector space, which deals with finite dimensional vectors. However, many vectors we encounter might have infinite dimensions (for example, most functions can be viewed as vectors have infinite dimensions), which we cannot analyze in vector spaces. That is why we are interested in Hilbert space. The concept of Hilbert space was put forward by David Hilbert in his work on quadratic forms in infinitely many variables. It's a generalization of Euclidean space to infinite dimensions. Due to its convenience in infinite dimensional vectors' analysis, Hilbert space has been widely used in other fields, for example physicians applied this concept in quantum mechanics, economists even used it in asset pricing theory.

In this paper, we give a brief introduction of Hilbert space, our paper is mainly based on Folland's book *Real Analysis:Modern Techniques and their Applications* (2nd edition) and Debnath and Mikusiński's book *Hilbert space with applications* (3rd edition).In second part, we first introduce the concept of inner product space, which is complex vector space equipped with inner product, and we also show that inner product space is a normed vector space with norm defined as a vector's inner product with itself. If an inner product space is complete, we call it a Hilbert space, which is showed in part 3. In part 4, we introduce orthogonal and orthonormal system and introduce the concept of orthonormal basis which is parallel to basis in linear vector space. In this part, we also give a brief introduction of orthogonal decomposition and Riesz representation theorem.

2 Inner Product Spaces

Definition 2.1(Inner product space)

Let E be a complex vector space. A mapping $\langle \, , \rangle \; E \times E \to C$.is called an *inner product* in E if for any $x, y, z \in E$ the following conditions are satisfied:

(a) $\langle x \, , y\rangle = \overline{\langle y \, , x\rangle}$(the bar denotes the complex conjugate);

(b) $\langle \alpha x + \beta y \, , z\rangle = \alpha\langle x \, , z\rangle + \beta\langle y \, , z\rangle$;

*Department of Economics, Cornell University, e-mail: yc2325@cornell.edu

†Department of Economics, Cornell University, e-mail: sl2555@cornell.edu

(c) $\langle x\,,x\rangle \geq 0$;

(d) $\langle x\,,x\rangle = 0$ implies $x = 0$.

Now let's look at some examples of inner product space:

Example 2.1 (N-dimensional inner product space) Consider space C^N of N-tuples $x = (x_1, ..., x_n)$ of complex numbers, with inner product defined by:

$$\langle x\,,y\rangle = \textstyle\sum_{k=1}^{N} x_k \overline{y_k}.$$

Example 2.2(infinite dimensional inner product space) Consider l^2 space of all sequences $\{x_n\}$ of complex numbers such that $\sum_{k=1}^{\infty} x_k < \infty$, the inner product is defined by:

$$\langle x\,,y\rangle = \textstyle\sum_{k=1}^{\infty} x_k \overline{y_k}.$$

Example 2.3(inner product space with functions) The space $\mathcal{C}([a,b])$ of all continuous complex valued functions on the interval $[a,b]$, with the inner product defined:

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

is an inner product space, this inner product we will use many times afterwards.

Inner product space is also called *pre-Hilbert space*. From the examples above, we can see that different from linear vector space, inner product space contains infinite dimensional vectors (such as functions), which is the reason why we want to study this space.

Next we want to show inner product space is also a normed vector space, with norm given by $||x|| = \langle x\,,x\rangle$. Before proving that, we first need to notice that Schwarz's inequality holds in inner product space.

Theorem 2.1(Schwarz's Inequality) For any two elements x and y of an inner product space,we have

$$|\langle x\,,y\rangle| \leq ||x||||y||$$

. The equality $|\langle x\,,y\rangle|$=$||x||||y||$ holds if and only if x and y are linearly dependent.

Theorem 2.2 Every inner product space is also a *normed vector space* with norm defined by $||x|| = \langle x\,,x\rangle$.

Proof:

1. $||x|| = 0$ if and only if $x = 0$;

2. $||\lambda x|| = \sqrt{\langle \lambda x\,,\lambda x\rangle} = \sqrt{\lambda\bar{\lambda}\langle x\,,x\rangle} = |\lambda|||x||$.

3. (Triangle Inequality) For any two elements x and y of an inner product space we have: $||x+y|| \leq ||x|| + ||y||$.

   This is because: $||x+y||^2 = \langle x+y\,,x+y\rangle = \langle x\,,x\rangle + 2Re\langle x\,,y\rangle + \langle y\,,y\rangle$

   $\leq \langle x\,,x\rangle + 2|\langle x\,,y\rangle| + \langle y\,,y\rangle$

   $\leq ||x||^2 + 2||x||||y|| + ||y||^2$ (From Schwarz's inequality)

   $= (||x|| + ||y||)^2$,

■

Next I give two theorems that are true in geometrics and they also hold in inner product space.

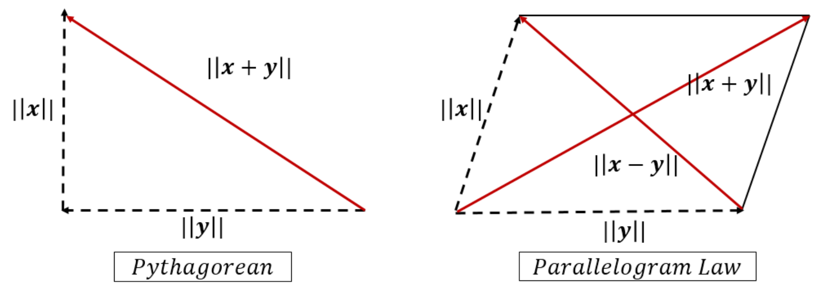

Figure 1: Pythagorean formula and Parallelogram Law

Theorem 2.3(Parallelogram Law) For any two elements x and y of an inner product space, we have

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$$

Theorem 2.4 (Pythagorean formula) For any pair of orthogonal vectors, we have

$$||x+y||^2 = ||x||^2 + ||y||^2$$

3 Hilbert Space

Definition 3.1(Hilbert space) A complete inner product space $\mathcal{H}$ is called a *Hilbert space.*

Now let's look at several examples:

Example 3.1 (Examples of Hilbert space)

(a) C is complete, it's Hilbert space, and so is $\mathcal{C}^N$.

(b) l^2 is a Hilbert space.

(c) $L^2(R)$ and $L^2([a,b])$ are Hilbert spaces.

Example 3.2 (Spaces that are not Hilbert spaces)

(a) Consider a space E consisting of sequences $\{x_n\}$ of complex numbers with only a finite number of nonzero terms.

$$x_n = (1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{n}, 0, 0, ...)$$

The inner product is defined as :

$$\langle x_n \, , x_m \rangle = \sum_{k=1}^{\infty} x_n \overline{x_m}$$

we can show that $\{x_n\}$ is a Cauchy sequence, since if $m > n$:

$$\lim_{m,n \to \infty} ||x_m - x_n|| = \lim_{m,n \to \infty} [\sum_{k=m}^{n} \frac{1}{k^2}]^{\frac{1}{2}} = 0$$

However, the sequence doesn't converge in E, since its limit has infinite terms that are not equal to 0.

(b) Consider the space $\mathcal{C}[0,1]$ consisting of continuous functions in [0,1]. Define

$$f_n(x) = \begin{cases} 1 & 0 \leq x \leq \frac{1}{2} \\ 1 - 2n(x - \frac{1}{2}) & \frac{1}{2} \leq x \leq \frac{1}{2n} + \frac{1}{2} \\ 0 & \frac{1}{2n} + \frac{1}{2} \leq x \leq 1 \end{cases}$$

Also it's easy to check $\{f_n(x)\}$ is Cauchy sequence,however,if we take the limit:

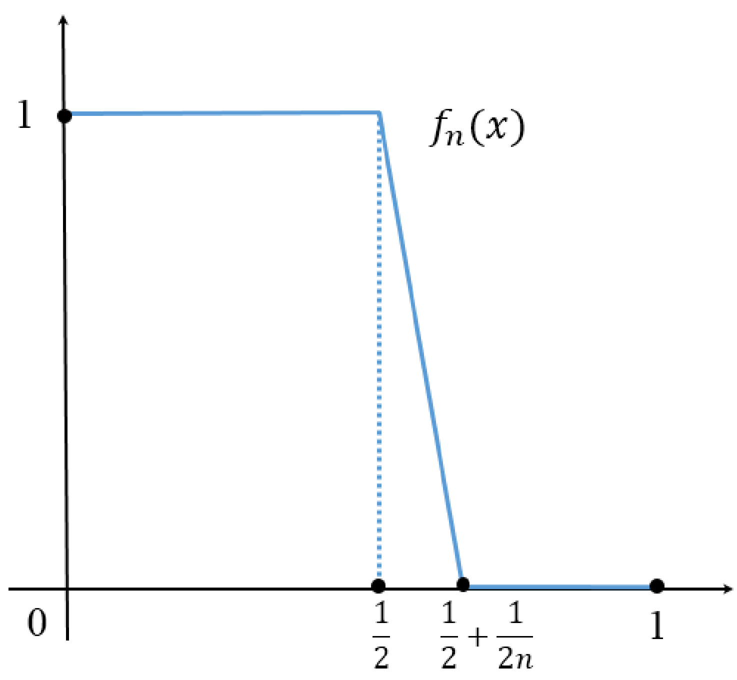

Figure 2: Example 3.2(b)

$$f(x) = \begin{cases} 1 & 0 \leq x \leq \frac{1}{2} \\ 0 & \frac{1}{2} < x \leq 1 \end{cases}$$

the limit function is not continuous, thus not in $\mathcal{C}[0,1]$, so $\mathcal{C}[0,1]$ is not Hilbert space.

4 Orthogonal and orthonormal system

In linear vector space E, we have *basis*, which is a linearly independent family $B = \{x_n, n = 1, 2, ...\}$ of vectors from E such that for any $x \in E$ x can be written as *finite* linear combination of x_n's: $x = \sum_{n=1}^{m} \lambda_n x_n$. As a result, we can represent whole elements in the space with simple "bricks"–basis. As we noted before, inner product space or Hilbert space can be viewed as a generalization of linear vector space. So in inner product space, we also expect to have such "bricks" to represent any element in the space. To achieve that goal, we will first introduce orthogonal and orthonormal system in this section, and then give the definition of orthonormal basis in inner product space, with this basis, we can represent any element in inner product space as *infinite* linear combination of elements in the basis.

4.1 Orthogonal and orthonormal system

Definition 4.1(Orthogonal and orthonormal system) In an inner product space E, suppose S is a family of nonzero vectors, such that if $x \neq y$, we have $x \perp y$ (i.e.$\langle x\,, y\rangle = 0$), we can S an *orthogonal system.* If for any $x \in S$, $||x|| = 1$, we can S an *orthonormal system.*

Why we want to introduce such system? Because elements in orthogonal system and orthonormal system preserve an important property as elements in a basis, which, as showed in the theorem below, is linearly independence.

Theorem 4.1(Linearly independence) Orthogonal systems are linearly independent.
Proof: Let S be an orthogonal system. Suppose $\sum_{k=1}^{n} \alpha_k x_k = 0$, for some $x_1, x_2, ..., x_n \in S$ and $\alpha_1, ..., \alpha_n \in C$.Then

$$0 = \sum_{m=1}^{n} \langle 0\,, \alpha_m x_m\rangle = \sum_{m=1}^{n} \langle \sum_{k=1}^{n} \alpha_k x_k\,, \alpha_m x_m\rangle = \sum_{m=1}^{n} |\alpha|^2 ||x_m||^2$$

This implies $\alpha_m = 0$. So $x_1, ..., x_n$ are linearly independent. ■
Example 4.1: For example, the function sequence $\{\phi_n(x)\}$, $\phi_n(x) = \frac{e^{i}nx}{\sqrt{2\pi}}$ is an orthonormal system in $L^2([-\pi, \pi])$. $\langle \phi_n\,, \phi_m\rangle = \int_{-\pi}^{\pi} \phi_n \overline{\phi_m} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)x} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos(n-m)x + i \sin(n-m)x) dx = 0$ for any $n \neq m$, thus if $\phi_n \neq \phi_m$, $\phi_n \perp \phi_m$. $\langle \phi_n\,, \phi_n\rangle = \int_{-\pi}^{\pi} \phi_n \overline{\phi_n} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-n)x} dx = 1$. So $||\phi_n|| = 1$ for any ϕ_n,$n \in N$.

Here we also have Pythagorean formula in orthogonal system for many vectors:

Theorem 4.2(Pythagorean Formula) Suppose $\{\, x_n\}$ is a sequence of orthogonal vectors in an inner product space. Then we have $|| \sum_{k=1}^{n} x_k||^2 = \sum_{k=1}^{n} ||x_k||^2$.

The proof of this theorem is very straight forward, we can use induction: first show this is true for $n = 2$ case (which is already given by Theorem 2.4). Then if $n = k - 1$ holds, using the fact that $\sum_{k=1}^{n} x_k = \sum_{k=1}^{n-1} x_k + x_n$, we can easily show this is also true for $n = k$.

Theorem 4.3 (Bessel's equality and inequality) Let $x_1, x_2, ..., x_n$ be an orthonormal set of vectors in inner product space E. For every $x \in E$,we have:

$$||x - \sum_{k=1}^{n} \langle x\,, x_k\rangle||^2 = ||x||^2 - \sum_{k=1}^{n} |\langle x\,, x_k\rangle|^2$$

and

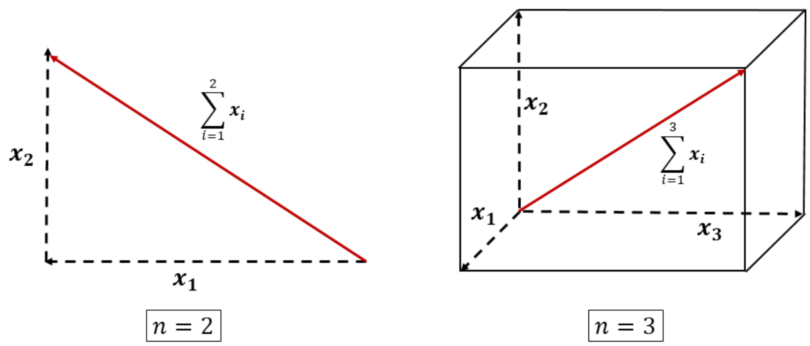

Figure 3: Pythagorean Theorem

$$\sum_{k=1}^{n} |\langle x\,, x_k\rangle|^2 \leq ||x||^2$$

Proof: using the result from Pythagorean theorem, we know: $|| \sum_{k=1}^{n} \alpha_k x_k||^2 = \sum_{k=1}^{n} ||\alpha_k x_k||^2 = \sum_{k=1}^{n} |\alpha_k|^2$.

For arbitrary complex numbers $\alpha_1, ..., \alpha_n$, we have: $||x - \sum_{k=1}^{n} \alpha_k x_k||^2 = \langle x - \sum_{k=1}^{n} \alpha_k x_k, x - \sum_{k=1}^{n} \alpha_k x_k\rangle = ||x||^2 - \sum_{k=1}^{n} |\langle x, x_k\rangle|^2 + \sum_{k=1}^{n} |\langle x, x_k\rangle - \alpha_k|^2$. Let $\alpha_k = \langle x, x_k\rangle$, we have $||x - \sum_{k=1}^{n}\langle x\,, x_k\rangle||^2 = ||x||^2 - \sum_{k=1}^{n} |\langle x\,, x_k\rangle|^2$. Since$||x - \sum_{k=1}^{n}\langle x\,, x_k\rangle||^2 \leq 0$, from Bessel's equality, we can easily know $\sum_{k=1}^{n} |\langle x\,, x_k\rangle|^2 \leq ||x||^2$. ■

4.2 Orthogonal decomposition and Riesz representation

From Bessel's inequality, we know $\sum_{k=1}^{n} |\langle x\,, x_k\rangle|^2 \leq ||x||^2$, if we let n to be ∞, we have $\sum_{k=1}^{\infty} |\langle x\,, x_k\rangle|^2 \leq ||x||^2$, thus $lim_{n\to\infty}|\langle x\,, x_n\rangle| = 0$, and for any n, let $f(x) = \langle x\,, x_n\rangle$, we can see $f(x)$ is a bounded linear function in inner product space E. In this section, we will introduce Riesz representation theorem, which claims that for any $x_0 \in E$, $f(x) = \langle x\,, x_0\rangle$ is a bounded linear function, besides any bounded linear function has such form.

Theorem 4.4(orthogonal decomposition) If $\mathcal{M}$ is a closed subspace of Hilbert space $\mathcal{H}$, then we have: $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$, where $\mathcal{M}^{\perp} = \{x \in \mathcal{H} : \langle x\,, y\rangle = 0, y \in \mathcal{M}\}$. In other words, for any $x \in \mathcal{H}$,x can be expressed uniquely as $x = y + z$ where $y \in \mathcal{M}$ and $z \in \mathcal{M}^{\perp}$. Moreover, y and z are the unique elements of $\mathcal{M}$ and $\mathcal{M}^{\perp}$ whose distance to x is smallest.

The proof of this theorem is a bit tedious and not of our interest, so I am gonna to skip that. One can easily find the proof in Folland's book(chapter 5).Now, I want to explain about the intuition behind that theorem.

As showed in the figure above, suppose the 3-dimension space as a whole is Hilbert space $\mathcal{H}$, x is an arbitrary vector in $\mathcal{H}$. Let $\mathcal{M}$ be a plane in $\mathcal{H}$, which is a closed set. Let y be the projection of x on $\mathcal{H}$, and let $z = x - y$, we have $z \perp y$, moreover, $z \perp$ any vector in $\mathcal{M}$, so $z \in \mathcal{M}^{\perp}$. Since y is projection of x, y is the element in $\mathcal{M}$ whose distance to x is smallest.

With Theorem 4.4, I can prove Riesz representation theorem.

Theorem 4.5 (Riesz Representation Theorem) If $f \in \mathcal{H}^*$,$\mathcal{H}^*$ is dual space of $\mathcal{H}$ (by

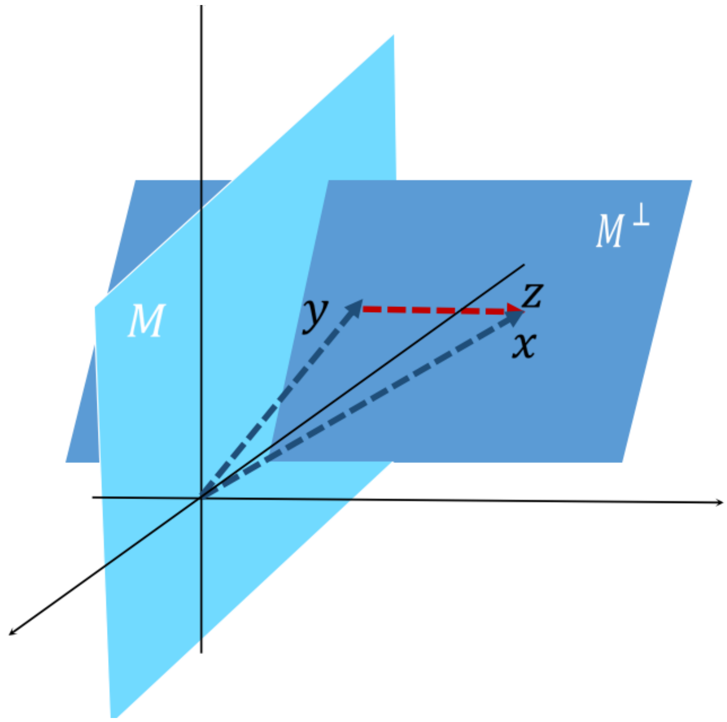

Figure 4: Understanding of orthogonal decomposition in Hilbert space

dual space we mean $\mathcal{H}^*$ consists of all bounded linear functions: $\mathcal{H} \to C$, then there is a unique $y \in \mathcal{H}$ such that $f(x) = \langle x\,, y\rangle$ for all $x \in E$.

Proof:

1.*Uniqueness*: If $f(x) = \langle x\,, y\rangle = \langle x\,, y'\rangle$ for all $x \in \mathcal{X}$. Take $x = y - y'$, we have $||y - y'||^2 = 0$, so $y = y'$.

2.*Existence*:

a) If $f(x) \equiv 0$, then we just pick $y = 0$, so we have $f(x) = \langle x\,, y\rangle$.

b) If $f(x)$ is not always equal to 0. Let $\mathcal{M} = x \in \mathcal{H} : f(x) = 0$. Since $f(x)$ is bounded linear function, $\mathcal{M}$ is a proper closed subset of E. By theorem 4.3, we know, $\mathcal{M}^\perp \neq \{0\}$, since if $\mathcal{M}^\perp = \{0\}$, for any $x \in \mathcal{H}$, we have $x = y + z, y \in \mathcal{M}, z \in \mathcal{M}^\perp$, so $x = y + 0 = y$,$f(x) = f(y) = 0$, for all x, contradiction! So we can pick a z in $\mathcal{M}^\perp$ such that $||z|| = 1$. Define $u = \underline{f(x)z} - f(z)x$, $\underline{\text{since}}$ $f(x)$ is linear function, we have $u \in \mathcal{M}$. So $0 = \langle u\,, z\rangle = f(x) - \langle x\,, \overline{f(z)}z\rangle$. Let $y = \overline{f(z)}z$, we have $f(x) = \langle x\,, y\rangle$. ■

4.3 Orthonormal Basis

Given an orthonormal sequence $\{x_n\}$, from Bessel's inequality, we have $\sum_{k=1}^{\infty} |\langle x\,, x_k\rangle|^2 \leq \infty$, thus $\{\langle x\,, x_n\rangle\} \in l^2$ so we have a mapping from inner product space E to l^2. The expansion

$$x \sim \sum_{n=1}^{\infty} \langle x\,, x_n\rangle x_n$$

is called a *generalized Fourier series* of x. The following theorem guarantees convergence of this series in Hilbert space.

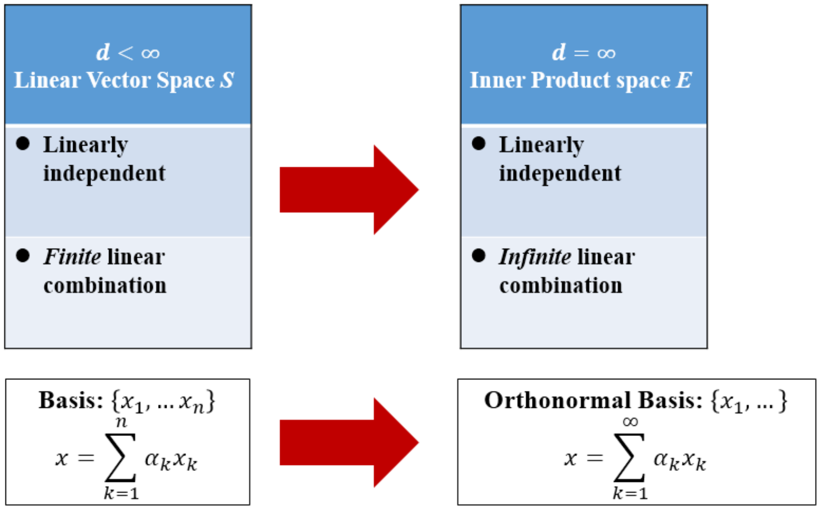

Figure 5: Extension of basis in inner product space

Theorem 4.6 Suppose $\{x_n\}$ is an orthonormal sequence in a Hilbert space $\mathcal{H}$, and $\{\alpha_n\}$ is a sequence of complex numbers. Then the series $\sum_{n=1}^{\infty} |\alpha|^2 < +\infty$ and in that case

$$|| \sum_{n=1}^{\infty} \alpha_n x_n||^2 = \sum_{n=1}^{\infty} |\alpha_n|^2$$

This can be viewed as a generalized case of Pythagorean theorem: the square of norm of infinite linear combinations of orthonormal sequence is equal to infinite sums of the coefficients' square. The proof is simple, we first consider the finite case (Pythagorean theorem) and use the completeness of $\mathcal{H}$ and get the result.

This theorem implies in a Hilbert space $\mathcal{H}$, the series $\sum_{n=1}^{\infty} \langle x , x_n \rangle x_n$ converges for every $x \in \mathcal{H}$, which means there exists a $x' \in \mathcal{H}$, such that $x' = \sum_{n=1}^{\infty} \langle x , x_n \rangle x_n$, however x' is not necessarily equal to x. If x' is equal to x, we call the sequence $\{x_n\}$ a *complete* orthonormal sequence.

Definition 4.2(Complete orthonormal sequence) An orthonormal sequence $\{x_n\}$ in an inner product space E is said to be *complete* if for every $x \in E$ we have

$$x = \sum_{n=1}^{\infty} \langle x , x_n \rangle x_n$$

In the beginning of this section, I recall the basic properties of our "bricks"– *basis* in linear vector space, and we also want to extend the properties of basis in infinite dimensional space. Here, we introduce *orthonormal basis.*

Definition 4.3(Orthonormal basis) An orthonormal system B is called an *orthonormal basis* if for every $x \in E$ has a *unique* representation:

$$x = \sum_{n=1}^{\infty} \alpha_n x_n$$

where $\alpha_n \in C$, and x_n's are distinct elements of B.

Note that a complete orthonormal sequence $\{x_n\}$ is an orthonormal basis. The proof is very straightforward.Suppose $x = \sum_{n=1}^{\infty} \alpha_n x_n = \sum_{n=1}^{\infty} \beta_n x_n$. We have $0 = ||x - x||^2 = ||\sum_{n=1}^{\infty}(\alpha_n - \beta_n)x_n||^2 = \sum_{n=1}^{\infty} |\alpha_n - \beta_n|^2$, so $\alpha_n = \beta_n$ for $b \in N$. Besides, if $\{x_n\}$ is an orthonormal basis in an inner product space, the span of $\{x_n\}$ (i.e. $span\{x_1, x_2, ...\} = \sum_{k=1}^{n} \alpha_k x_k$)is dense in E.

One famous example of orthonormal basis is given below:

Example 4.2 The sequence of functions: $\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, ...$ is a complete orthonormal system in $L^2([-\pi, \pi])$, so for any function in $L^2([-\pi, \pi])$, we can express $f(x)$ as infinite linear combination of the series above, which is the result given by Fourier transformation.