A Brief Introduction to Hilbert Space Theory with RKHS in Mind

Aydın Aytuna

2024

These notes are written for the ”RKHS Learning Seminars” at the Institute of Applied Mathematics, METU. It aims to introduce the audience to the fundamental ideas of elementary Hilbert space theory. We assume the participants have good knowledge of linear algebra and advanced calculus. The material covered is relatively standard and contains no new mathematics. The book ”A Primer on Reproducing Kernel Hilbert Spaces.” by J.H. Manton and P. Amblard was chosen as the principal reference for this seminar series. Hence, it shaped the structure of these notes. Additional references used in these notes plus references that can be used for further studies are listed in bibliography.

Let V be a vector space over $\mathbb{C}$ (or $\mathbb{R}$). Let $B : V \times V \to \mathbb{C}$ be a function satisfying

$$\begin{aligned} &\bullet B(x,y) = \overline{B(y,x)} \quad \forall x,y \in V, \\ &\bullet B(x+\lambda y,\xi) = B(x,\xi) + \lambda B(y,\xi) \quad \forall x,y,\xi \in V, \lambda \in \mathbb{C}, \\ &\bullet B(x,x) \geq 0 \quad \forall x \in V. \end{aligned} \tag{*}$$

An important property of such functions (usually refered to as *sesquilinear fuctions*) is:

$$|B(x,y)| \leq B(x,x)^{1/2} B(y,y)^{1/2} \quad \forall x,y \in V. \tag{1}$$

To see this let us look at:

$$\begin{aligned} 0 \leq B(x+ty,x+ty) &= B(x,x) + tB(y,x) + \bar{t}B(x,y) + |t|^2 B(y,y) \\ &= B(x,x) + 2\Re\mathfrak{e}(tB(x,y)) + |t|^2 B(y,y) \qquad \forall x,y \in V, t \in \mathbb{C} \end{aligned}$$

Choosing $t = sB(y,x)$, $s \in \mathbb{R}$, we have;

$$0 \leq B(x,x) + 2s \mid B(x,y) \mid^2 + s^2 \mid B(y,x) \mid^2 B(y,y).$$

So the discriminant of this polynomial $P(s)$ must satisfy,

$$4B(x,y)^4 - 4B(x,x)B(y,y) \mid B(y,x) \mid^2 \leq 0$$

or

$$\begin{aligned} &\mid B(x,y) \mid^4 \leq B(x,x)B(y,y) \mid B(y,x) \mid^2 \\ &\mid B(x,y) \mid^2 \leq B(x,x)B(y,y). \end{aligned}$$

Let V be a vector space over $\mathbb{C}$ (or $\mathbb{R}$). Let $B : V \times V \to \mathbb{C}$ be a function satisfying

$$\begin{aligned} &\bullet B(x,y) = \overline{B(y,x)} \quad \forall x,y \in V, \\ &\bullet B(x+\lambda y,\xi) = B(x,\xi) + \lambda B(y,\xi) \quad \forall x,y,\xi \in V, \lambda \in \mathbb{C}, \\ &\bullet B(x,x) \geq 0 \quad \forall x \in V. \end{aligned} \tag{*}$$

An important property of such functions (usually refered to as *sesquilinear fuctions*) is:

$$|B(x,y)| \leq B(x,x)^{1/2} B(y,y)^{1/2} \quad \forall x,y \in V. \tag{1}$$

To see this let us look at:

$$\begin{aligned} 0 \leq B(x+ty,x+ty) &= B(x,x) + tB(y,x) + \bar{t}B(x,y) + |t|^2 B(y,y) \\ &= B(x,x) + 2\mathfrak{Re}(tB(x,y)) + |t|^2 B(y,y) \qquad \forall x,y \in V, t \in \mathbb{C} \end{aligned}$$

Choosing $t = sB(y,x)$, $s \in \mathbb{R}$, we have;

$$0 \leq B(x,x) + 2s \mid B(x,y) \mid^2 + s^2 \mid B(y,x) \mid^2 B(y,y).$$

So the discriminant of this polynomial $P(s)$ must satisfy,

$$4B(x,y)^4 - 4B(x,x)B(y,y) \mid B(y,x) \mid^2 \leq 0$$

or

$$\begin{aligned} &\mid B(x,y) \mid^4 \leq B(x,x)B(y,y) \mid B(y,x) \mid^2 \\ &\mid B(x,y) \mid^2 \leq B(x,x)B(y,y). \end{aligned}$$

This simple observation is the source of many inequalities, such as

$$\Big|\sum_i x_i\overline{y_i}\Big| \leq \Big(\sum |\, x_i \,|^2\Big)^{1/2}\Big(\sum |\, y_i \,|^2\Big)^{1/2},$$

or its continuous analog

$$\Big|\int_a^b f(t)\overline{g(t)}dt\Big| \leq \Big(\int_a^b |\, f(t) \,|^2 \, dt\Big)^{1/2}\Big(\int_a^b |\, g(t) \,|^2 \, dt\Big)^{1/2} \quad \forall f, g \in C[a, b].$$

Another property of such functions is:

$$\begin{aligned} B(x+y, x+y) &\leq B(x,x) + 2\Re\mathfrak{e}B(x,y) + B(y,y) \\ &\leq B(x,x) + 2B(x,x)^{1/2}B(y,y)^{1/2} + B(y,y) \\ &= \left(B(x,x)^{1/2} + B(y,y)^{1/2}\right)^2. \end{aligned} \quad (2)$$

Recall that

An **inner product** $\langle\, , \rangle$ on $(V, \mathbb{C})$ is a function on $V \times V$ that satisfy the conditions (*) with the additional condition:

$$\langle x, x\rangle = 0 \iff x = 0.$$

This simple observation is the source of many inequalities, such as

$$\Big|\sum_i x_i \overline{y_i}\Big| \leq \Big(\sum | x_i |^2 \Big)^{1/2} \Big(\sum | y_i |^2 \Big)^{1/2},$$

or its continuous analog

$$\Big|\int_a^b f(t)\overline{g(t)}dt\Big| \leq \Big(\int_a^b | f(t) |^2 \, dt\Big)^{1/2} \Big(\int_a^b | g(t) |^2 \, dt\Big)^{1/2} \quad \forall f, g \in C[a, b].$$

Another property of such functions is:

$$\begin{aligned} B(x+y, x+y) &\leq B(x,x) + 2\Re\mathfrak{e} B(x,y) + B(y,y) \\ &\leq B(x,x) + 2B(x,x)^{1/2} B(y,y)^{1/2} + B(y,y) \\ &= \left(B(x,x)^{1/2} + B(y,y)^{1/2}\right)^2. \end{aligned} \tag{2}$$

Recall that

An **inner product** $\langle \, , \rangle$ on $(V, \mathbb{C})$ is a function on $V \times V$ that satisfy the conditions (*) with the additional condition:

$$\langle x, x \rangle = 0 \iff x = 0.$$

An inner product gives a "norm" on V by setting:

$$\| x \| = \langle x, x \rangle^{1/2},$$

which can be thought as a "distance of x to the vector 0. In fact

$$d(x, y) = \| x - y \|$$

satisfies the usual conditions of a metric on V. Namely:

- $d(x, y) \leq d(x, z) + d(z, y)$,
- $d(x, y) = d(y, x)$,
- $d(x, x) \geq 0$ and $d(x, x) = 0 \Leftrightarrow x = 0$.

Hence an inner product on V induces a notion of **"distance"** between points in V, therefore introduce notions like **"convergence"** of sequences via

$$x_n \to x \Leftrightarrow \| x_n - x \| \to 0$$

and **"continuity"** of functions $f : V_1 \to V_2$ between inner product spaces via

$$f \text{ is } \textbf{continuous} \text{ at a point } x \text{ in case: } x_n \to x \in V_1 \implies f(x_n) \to f(x) \in V_2.$$

Note that for linear operators convergence can be checked by using norms.

An inner product gives a "norm" on V by setting:

$$\| x \| = \langle x, x \rangle^{1/2},$$

which can be thought as a "distance of x to the vector 0. In fact

$$d(x, y) = \| x - y \|$$

satisfies the usual conditions of a metric on V. Namely:

- $d(x, y) \leq d(x, z) + d(z, y)$,
- $d(x, y) = d(y, x)$,
- $d(x, x) \geq 0$ and $d(x, x) = 0 \Leftrightarrow x = 0$.

Hence an inner product on V induces a notion of **"distance"** between points in V, therefore introduce notions like **"convergence"** of sequences via

$$x_n \to x \Leftrightarrow \| x_n - x \| \to 0$$

and **"continuity"** of functions $f : V_1 \to V_2$ between inner product spaces via

$$f \text{ is } \textbf{continuous} \text{ at a point } x \text{ in case: } x_n \to x \in V_1 \implies f(x_n) \to f(x) \in V_2.$$

Note that for linear operators convergence can be checked by using norms.

Theorem

For a linear operator $T : (V_1, \langle\,,\rangle_1) \to (V_2, \langle\,,\rangle_2)$ *the following are equivalent:*

(i) *T is continuous at the point* 0*,*

(ii) *T is continuous at every point of V* (3)

(iii) $\exists c > 0\, ;\ \| T(x) \|_2 \leq c \| x \|_1$.

Proof.

To see "$(i) \implies (ii)$", suppose T is continuous at 0. Fix x and suppose $x_n \to x$, i.e. $x_n - x \to 0$, then $T(x_n) - T(x) = T(x_n - x) \to T(0) = 0$ or, T is continuous at x.

To show "$(i) \implies (iii)$", suppose no such c exists. Then for each $N \in \mathbb{N}$, we can find $x_N \neq 0$ such that

$$\| T(x_N) \|_2 \geq N \| x_N \|_1 .$$

Consider the sequence $\{u_N\}_N$, where $u_N := \frac{x_N}{\sqrt{N}\|x_N\|_1}$. We have:

$$\| u_N \|_1 = \left\| \frac{x_N}{\sqrt{N} \| x_N \|_1} \right\|_1 = \frac{\| x_N \|_1}{\sqrt{N} \| x_N \|_1} = \frac{1}{\sqrt{N}} \to 0 \quad \text{as } N \to \infty.$$

This implies $T(u_N) \to 0$, but we also have:

$$\| T(u_N) \|_2 = \frac{\| T(x_N) \|_2}{\sqrt{N} \| x_N \|_1} \geq \frac{N \| x_N \|_1}{\sqrt{N} \| x_N \|_1} = \sqrt{N},$$

so the assumption leads to a contradiction. Rest of the assertion is self evident. □

Theorem

For a linear operator $T : (V_1, \langle\,,\rangle_1) \to (V_2, \langle\,,\rangle_2)$ *the following are equivalent:*

(i) *T is continuous at the point* 0*,*

(ii) *T is continuous at every point of V* (3)

(iii) $\exists c > 0\,;\ \| T(x) \|_2 \leq c \| x \|_1$.

Proof.

To see "$(i) \implies (ii)$", suppose T is continuous at 0. Fix x and suppose $x_n \to x$, i.e. $x_n - x \to 0$, then $T(x_n) - T(x) = T(x_n - x) \to T(0) = 0$ or, T is continuous at x.

To show "$(i) \implies (iii)$", suppose no such c exists. Then for each $N \in \mathbb{N}$, we can find $x_N \neq 0$ such that

$$\| T(x_N) \|_2 \geq N \| x_N \|_1 \ .$$

Consider the sequence $\{u_N\}_N$, where $u_N := \frac{x_N}{\sqrt{N}\|x_N\|_1}$. We have:

$$\| u_N \|_1 = \left\| \frac{x_N}{\sqrt{N} \| x_N \|_1} \right\|_1 = \frac{\| x_N \|_1}{\sqrt{N} \| x_N \|_1} = \frac{1}{\sqrt{N}} \to 0 \quad \text{as } N \to \infty.$$

This implies $T(u_N) \to 0$, but we also have:

$$\| T(u_N) \|_2 = \frac{\| T(x_N) \|_2}{\sqrt{N} \| x_N \|_1} \geq \frac{N \| x_N \|_1}{\sqrt{N} \| x_N \|_1} = \sqrt{N},$$

so the assumption leads to a contradiction. Rest of the assertion is self evident. □

Remark

In the special case of finite dimensional inner product spaces V_1 and V_2, every operator $T : V_1 \to V_2$ is automatically continuous since if $\{e_1, \cdots, e_N\}$ is an orthonormal basis in a finite dimensional inner product space $(V_1, \langle\, , \rangle)$, then, if $x_n \to x$, since

$$x_n = c_1^n e_1 + \cdots + c_N^n e_N = \sum_{i=1}^{N} c_i^n e_i, \quad x = c_1 e_1 + \cdots + c_N e_N = \sum_{i=1}^{N} c_i e_i$$

we have:

$$\begin{aligned} \| x_n - x \|^2 &= \left\langle \textstyle\sum_{i=1}^N c_i^n e_i - \sum_{i=1}^N c_i e_i, \sum_{i=1}^N c_i^n e_i - \sum_{i=1}^N c_i e_i \right\rangle = \left\langle \textstyle\sum_{i=1}^N (c_i^n - c_i) e_i, \sum_{i=1}^N (c_i^n - c_i) e_i \right\rangle \\ &= \textstyle\sum_{i=1}^N (c_i^n - c_i) \cdot \left\langle e_i, \sum_{j=1}^N (c_j^n - c_j) e_j \right\rangle = \sum_{i=1}^N \sum_{j=1}^N (c_i^n - c_i) \overline{(c_j^n - c_j)} \langle e_i, e_j \rangle \\ &= \textstyle\sum_{i=1}^N |c_i^n - c_i|^2, \end{aligned}$$

since $\langle e_i, e_j \rangle = 0$ for $i \neq j$ and $\langle e_i, e_i \rangle = \| e_i \|^2 = 1$. So $x_n - x \to 0 \iff | c_i^n - c_i | \to 0$, $\forall i = 1, \cdots, N$, $\iff c_i^n \to c_i$ $\forall i = 1, \cdots, N$.

With respect to orthonormal basis $\{e_i\}_{i=1}^N$ of V_1 and $\{f_j\}_{j=1}^M$ of V_2, T can be represented as a matrix $A = \{a_{ij}\}$, where $Te_i = \sum_{j=1}^M a_{ij} f_j$, so

$$T(x) = T\Big(\sum_{i=1}^N x_i e_i\Big) = \sum_{j=1}^M \Big(\sum_{i=1}^N a_{ij} x_i\Big) f_j, \quad T(x_n - x) = \sum_{j=1}^M \sum_{i=1}^N a_{ij} (x_i^n - x_i) f_j.$$

Hence $Tx_n \to Tx$ as $x_n \to x$.

Remark

In the special case of finite dimensional inner product spaces V_1 and V_2, every operator $T : V_1 \to V_2$ is automatically continuous since if $\{e_1, \cdots, e_N\}$ is an orthonormal basis in a finite dimensional inner product space $(V_1, \langle\, , \rangle)$, then, if $x_n \to x$, since

$$x_n = c_1^n e_1 + \cdots + c_N^n e_N = \sum_{i=1}^{N} c_i^n e_i, \quad x = c_1 e_1 + \cdots + c_N e_N = \sum_{i=1}^{N} c_i e_i$$

we have:

$$\begin{aligned} \| x_n - x \|^2 &= \left\langle \textstyle\sum_{i=1}^N c_i^n e_i - \sum_{i=1}^N c_i e_i, \sum_{i=1}^N c_i^n e_i - \sum_{i=1}^N c_i e_i \right\rangle = \left\langle \textstyle\sum_{i=1}^N (c_i^n - c_i) e_i, \sum_{i=1}^N (c_i^n - c_i) e_i \right\rangle \\ &= \textstyle\sum_{i=1}^N (c_i^n - c_i) \cdot \left\langle e_i, \sum_{j=1}^N (c_j^n - c_j) e_j \right\rangle = \sum_{i=1}^N \sum_{j=1}^N (c_i^n - c_i) \overline{(c_j^n - c_j)} \langle e_i, e_j \rangle \\ &= \textstyle\sum_{i=1}^N |c_i^n - c_i|^2, \end{aligned}$$

since $\langle e_i, e_j \rangle = 0$ for $i \neq j$ and $\langle e_i, e_i \rangle = \| e_i \|^2 = 1$. So $x_n - x \to 0 \iff | c_i^n - c_i | \to 0, \forall i = 1, \cdots, N,$ $\iff c_i^n \to c_i \, \forall i = 1, \cdots, N.$

With respect to orthonormal basis $\{e_i\}_{i=1}^N$ of V_1 and $\{f_j\}_{j=1}^M$ of V_2, T can be represented as a matrix $A = \{a_{ij}\}$, where $Te_i = \sum_{j=1}^M a_{ij} f_j$, so

$$T(x) = T\Big(\sum_{i=1}^N x_i e_i\Big) = \sum_{j=1}^M \Big(\sum_{i=1}^N a_{ij} x_i\Big) f_j, \quad T(x_n - x) = \sum_{j=1}^M \sum_{i=1}^N a_{ij}(x_i^n - x_i) f_j.$$

Hence $Tx_n \to Tx$ as $x_n \to x$.

At this point, we would like to recall **Gram-Schmidt orthogonalization process**:

Given n linearly independent elements $v_1, \cdots, v_n$ in an inner product space $(V, \langle\, , \rangle)$, consider the inductively defined sequence of vectors:

$$u_1 = v_1, \quad u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1, \quad \cdots, \quad u_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, u_i \rangle}{\langle u_i, u_i \rangle} u_i.$$

A routine check will show that u_i's for $1 \leq k \leq n$, are orthogonal to each other, i.e, $\langle u_i, u_j \rangle = 0$, and the normalized $\left\{ \frac{u_i}{\|u_i\|} \right\}$ vectors form an orthonormal basis for span$\{v_1, \cdots, v_n\}$.

Also observe that each $u_k \in$ span$\{v_1, \cdots, v_k\}$.

This algorithm works also in infinite dimensional inner product spaces if a sequence of finitely linearly independent vectors $\{v_i\}_{i=1}^{\infty}$ are given, and it yields a sequence orthonormal vectors $\{u_i\}_{i=1}^{\infty}$ spanning the vector subspace spanned by $\{v_i\}$'s.

Automatic continuity of linear operators fails if the dimension is not finite.

At this point, we would like to recall **Gram-Schmidt orthogonalization process**:

Given n linearly independent elements $v_1, \cdots, v_n$ in an inner product space $(V, \langle\, , \rangle)$, consider the inductively defined sequence of vectors:

$$u_1 = v_1, \quad u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle} u_1, \quad \cdots, \quad u_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, u_i\rangle}{\langle u_i, u_i\rangle} u_i.$$

A routine check will show that u_i's for $1 \le k \le n$, are orthogonal to each other, i.e, $\langle u_i, u_j\rangle = 0$, and the normalized $\left\{\frac{u_i}{\|u_i\|}\right\}$ vectors form an orthonormal basis for span$\{v_1, \cdots, v_n\}$.

Also observe that each $u_k \in$ span$\{v_1, \cdots, v_k\}$.

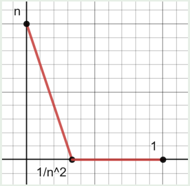
This algorithm works also in infinite dimensional inner product spaces if a sequence of finitely linearly independent vectors $\{v_i\}_{i=1}^{\infty}$ are given, and it yields a sequence orthonormal vectors $\{u_i\}_{i=1}^{\infty}$ spanning the vector subspace spanned by $\{v_i\}$'s.

Automatic continuity of linear operators fails if the dimension is not finite.

Example 0

Let $V = (C[0,1], \langle\, , \rangle)$ be the vector space of real valued continuous functions on $[0,1]$ with the inner product

$$\langle f, g\rangle := \int_0^1 f(t)g(t)dt.$$



Let $\Bbbk_0 : C[0,1] \to (\mathbb{R}, \langle\, , \rangle)$ be the linear operator $\Bbbk_0(f) := f(0)$.

Consider the sequence $\{f_n\}$ in $C[0,1]$ defined as:

$$f_n(t) = \begin{cases} -n^3 t + n & 0 \leq t \leq \frac{1}{n^2}, \\ 0 & \frac{1}{n^2} < t \leq 1. \end{cases}$$

Note $\| f_n \| = \int_0^1 f_n(t)^2 dt \leq 1$, yet $\| \Bbbk_0(f_n) \| = | f_n(0) | = n$, so $\exists$ no $C > 0$ such that

$$\| \Bbbk_0(f) \| \leq C \| f \| \quad \forall f \in C[0,1].$$

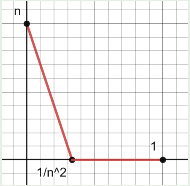
Hence $\Bbbk_0$ is not continuous.

This is one of the reasons why continuity (in general the topology) is suppressed in finite dimensional linear algebra and why in infinite dimensional linear algebra topology plays an important role if one wants to develop a reasonable theory.

Example 0

Let $V = (C[0,1], \langle\, , \rangle)$ be the vector space of real valued continuous functions on $[0,1]$ with the inner product

$$\langle f, g\rangle := \int_0^1 f(t)g(t)dt.$$



Let $\mathbb{k}_0 : C[0,1] \to (\mathbb{R}, \langle\, , \rangle)$ be the linear operator $\mathbb{k}_0(f) := f(0)$.

Consider the sequence $\{f_n\}$ in $C[0,1]$ defined as:

$$f_n(t) = \begin{cases} -n^3 t + n & 0 \le t \le \frac{1}{n^2}, \\ 0 & \frac{1}{n^2} < t \le 1. \end{cases}$$

Note $\| f_n \| = \int_0^1 f_n(t)^2 dt \le 1$, yet $\| \mathbb{k}_0(f_n) \| = | f_n(0) | = n$, so $\exists$ no $C > 0$ such that

$$\| \mathbb{k}_0(f) \| \le C \| f \| \quad \forall f \in C[0,1].$$

Hence $\mathbb{k}_0$ is not continuous.

This is one of the reasons why continuity (in general the topology) is suppressed in finite dimensional linear algebra and why in infinite dimensional linear algebra topology plays an important role if one wants to develop a reasonable theory.

We close this section with an elementary observation about inner product spaces that generalizes the **Parallelogram law** of elementary geometry:

$$\begin{aligned} \| x - y \|^2 + \| x + y \|^2 &= \langle x - y, x - y\rangle + \langle x + y, x + y\rangle \\ &= \langle x, x\rangle - 2\Re\mathfrak{e}\langle x, y\rangle + \langle y, y\rangle + \langle x, x\rangle + 2\Re\mathfrak{e}\langle x, y\rangle + \langle y, y\rangle \\ &= 2 \| x \|^2 + 2 \| y \|^2, \end{aligned}$$

or

$$\| x + y \|^2 + \| x - y \|^2 = 2 \| x \|^2 + 2 \| y \|^2 .$$

If a sequence "clusters" in an inner product space, it is desirable that it converges somewhere. In other words if the distances between the points of the sequence gets very small as one proceeds to the tail of the sequence (in some sense one wants to assume that there are no "holes" in the space!)

Example

In $C[0,1]$ of **Example 0**, for $n \geq 2$ let:

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

$$\| f_n - f_m \|^2 = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} | (f_n - f_m)(t) |^2 \, dt \leq \frac{4}{n} \text{ if } m > n.$$

So $\lim_{n,m\to\infty} \| f_n - f_m \|^2 = 0$, and yet if $\exists F \in C[0,1]$ such that $\| f_n - F \|^2 \to 0$;

$$\| f_n - F \|^2 = \int_0^{\frac{1}{2}-\frac{1}{n}} | F - 1 |^2 + \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} | F - f_n |^2 + \int_{\frac{1}{2}}^{1} | F |^2 \to 0.$$

Then $F \equiv 0$ on $[1/2, 1]$ and $F \equiv 1$ on any closed interval $[0, \alpha]$, $\alpha < 1/2$. This means:

$$1 = \lim_{t\to(1/2)^-} F(t) = \lim_{t\to(1/2)^+} F(t) = 0.$$

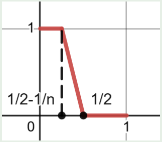
So $\exists$ no such $F \in C[0,1]$.

If a sequence "clusters" in an inner product space, it is desirable that it converges somewhere. In other words if the distances between the points of the sequence gets very small as one proceeds to the tail of the sequence (in some sense one wants to assume that there are no "holes" in the space!)

Example

In $C[0,1]$ of **Example 0**, for $n \geq 2$ let:

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



$$\| f_n - f_m \|^2 = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} | (f_n - f_m)(t) |^2 \, dt \leq \frac{4}{n} \text{ if } m > n.$$

So $\lim_{n,m\to\infty} \| f_n - f_m \|^2 = 0$, and yet if $\exists F \in C[0,1]$ such that $\| f_n - F \|^2 \to 0$;

$$\| f_n - F \|^2 = \int_0^{\frac{1}{2}-\frac{1}{n}} | F - 1 |^2 + \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} | F - f_n |^2 + \int_{\frac{1}{2}}^{1} | F |^2 \to 0.$$

Then $F \equiv 0$ on $[1/2, 1]$ and $F \equiv 1$ on any closed interval $[0, \alpha]$, $\alpha < 1/2$. This means:

$$1 = \lim_{t\to(1/2)^-} F(t) = \lim_{t\to(1/2)^+} F(t) = 0.$$

So $\exists$ no such $F \in C[0,1]$.

Let us establish some terminology:

Definition

- A sequence $\{x_n\}$ in an inner product space is said to be **Cauchy** in case $\forall \varepsilon > 0\ \exists N$ such that if $n, m \geq N$ then $\| x_n - x_m \| \leq \varepsilon$
- An inner product space $(V, \langle\, , \rangle)$ is called **complete** in case every Cauchy sequence converges to a point in V
- A complete inner product space is called a **Hilbert space**.

Main Example

$$\ell_2 = \{(x_n) : \sum_n | x_n |^2 < \infty\} \text{ with } \langle x, y\rangle := \sum_n x_n \overline{y_n}, \text{ for } x = (x_n), y = (y_n).$$

This makes sense by Equation (1) of Part 1. If $x^\alpha = \{x_n^\alpha\}_{n=1}^\infty, \alpha = 1, 2, \cdots$ is a Cauchy sequence in ℓ_2 then fix $\varepsilon > 0$, $\exists M_\varepsilon$ for every N

$$\| x^\alpha - x^\beta \|^2 = \sum_{n=1}^{N} | x_n^\alpha - x_n^\beta |^2 + \sum_{n=N+1}^{\infty} | x_n^\alpha - x_n^\beta |^2 \leq \varepsilon^2/4 \quad \text{if } \alpha, \beta \geq M.$$

Then $\xi_N^\alpha := \{x_n^\alpha\}_{n=1}^N$ is Cauchy in $\mathbb{C}^N$, so it converges in $\mathbb{C}^N$.

In particular, $\exists x = (x_n)_{n=1}^\infty$ such that $x_n^\alpha \xrightarrow{\alpha\to\infty} x_n\ \forall n$ and $\xi_N^\alpha \to (x_1, \cdots, x_N)$ in $\mathbb{C}^N$.

Let us establish some terminology:

Definition

- A sequence $\{x_n\}$ in an inner product space is said to be **Cauchy** in case $\forall \varepsilon > 0\ \exists N$ such that if $n, m \geq N$ then $\| x_n - x_m \| \leq \varepsilon$
- An inner product space $(V, \langle\,,\rangle)$ is called **complete** in case every Cauchy sequence converges to a point in V
- A complete inner product space is called a **Hilbert space**.

Main Example

$$\ell_2 = \{(x_n) : \sum_n | x_n |^2 < \infty\} \text{ with } \langle x, y\rangle := \sum_n x_n \overline{y_n}, \text{ for } x = (x_n), y = (y_n).$$

This makes sense by Equation (1) of Part 1. If $x^\alpha = \{x_n^\alpha\}_{n=1}^\infty, \alpha = 1, 2, \cdots$ is a Cauchy sequence in ℓ_2 then fix $\varepsilon > 0$, $\exists M_\varepsilon$ for every N

$$\| x^\alpha - x^\beta \|^2 = \sum_{n=1}^{N} | x_n^\alpha - x_n^\beta |^2 + \sum_{n=N+1}^{\infty} | x_n^\alpha - x_n^\beta |^2 \leq \varepsilon^2/4 \quad \text{if } \alpha, \beta \geq M.$$

Then $\xi_N^\alpha := \{x_n^\alpha\}_{n=1}^N$ is Cauchy in $\mathbb{C}^N$, so it converges in $\mathbb{C}^N$.

In particular, $\exists x = (x_n)_{n=1}^\infty$ such that $x_n^\alpha \xrightarrow{\alpha\to\infty} x_n\ \forall n$ and $\xi_N^\alpha \to (x_1, \cdots, x_N)$ in $\mathbb{C}^N$.

Main Example(Cont.)

Therefore for any S, and $\alpha \geq M_\varepsilon$, choose β so that $\beta \geq M_\varepsilon$ and $\sum_{n=1}^{S} \mid x_n^\beta - x_n \mid^2 \leq \varepsilon^2/4$.

$$\Longrightarrow \quad \left(\sum_{n=1}^{S} \mid x_n^\alpha - x_n \mid^2\right)^{1/2} \leq \left(\sum_{n=1}^{S} \mid x_n^\alpha - x_n^\beta \mid^2\right)^{1/2} + \left(\sum_{n=1}^{S} \mid x_n^\beta - x_n \mid^2\right)^{1/2} \leq \varepsilon,$$

$$\Longrightarrow \quad \sum_{n=1}^{\infty} \mid x_n^\alpha - x_n \mid^2 \leq \varepsilon \quad \forall \alpha \geq M_\varepsilon \implies (x_n) \in \ell_2 \text{ and } x^\alpha \to x.$$

So ℓ_2 is a Hilbert space.

Notation

For any subset S of a Hilbert space $(H, \langle \, , \rangle)$, $S^\perp$ will denote the set of all elements of H that are orthogonal to elements of S. That is:

$$S^\perp := \{x \in H : \langle x, s \rangle = 0 \, \forall s \in S\}.$$

Clearly $S^\perp$ is a subspace, and is closed in the sense that it contains the limit points of all Cauchy sequences in it, since $x_n \in S^\perp$, $n = 1, 2, \cdots$, and $x_n \to x$ implies that

$$\mid \langle x, s \rangle \mid = \mid \langle x_n, s \rangle - \langle x, s \rangle \mid = \mid \langle x_n - x, s \rangle \mid \leq \| x_n - x \| \| s \| \xrightarrow{n \to \infty} 0 \quad \Longrightarrow x \in S^\perp.$$

Hence $S^\perp$ is itself a Hilbert space under $\langle \, , \rangle$.

Main Example(Cont.)

Therefore for any S, and $\alpha \geq M_\varepsilon$, choose β so that $\beta \geq M_\varepsilon$ and $\sum_{n=1}^{S} | x_n^\beta - x_n |^2 \leq \varepsilon^2/4$.

$$\Longrightarrow \quad \left(\sum_{n=1}^{S} | x_n^\alpha - x_n |^2 \right)^{1/2} \leq \left(\sum_{n=1}^{S} | x_n^\alpha - x_n^\beta |^2 \right)^{1/2} + \left(\sum_{n=1}^{S} | x_n^\beta - x_n |^2 \right)^{1/2} \leq \varepsilon,$$

$$\Longrightarrow \quad \sum_{n=1}^{\infty} | x_n^\alpha - x_n |^2 \leq \varepsilon \quad \forall \alpha \geq M_\varepsilon \implies (x_n) \in \ell_2 \text{ and } x^\alpha \to x.$$

So ℓ_2 is a Hilbert space.

Notation

For any subset S of a Hilbert space $(H, \langle\, , \rangle)$, $S^\perp$ will denote the set of all elements of H that are orthogonal to elements of S. That is:

$$S^\perp := \{x \in H : \langle x, s \rangle = 0 \, \forall s \in S\}.$$

Clearly $S^\perp$ is a subspace, and is closed in the sense that it contains the limit points of all Cauchy sequences in it, since $x_n \in S^\perp$, $n = 1, 2, \cdots$, and $x_n \to x$ implies that

$$| \langle x, s \rangle | = | \langle x_n, s \rangle - \langle x, s \rangle | = | \langle x_n - x, s \rangle | \leq \| x_n - x \| \| s \| \xrightarrow{n \to \infty} 0 \quad \Longrightarrow x \in S^\perp.$$

Hence $S^\perp$ is itself a Hilbert space under $\langle\, , \rangle$.

In practice, one needs to know the existence of a point in a given set $C \not\ni 0$ of a Hilbert space that is closest to 0, and if it exists, whether this point is unique.

One can get a satisfactory answer if C is a **closed convex** set in H. Closed in the sense that it contains all its limit points (i.e., $x_n \to x, x_n \in C \implies x \in C$), and convex in the sense that $\forall x, y \in C$, the midpoint $(x+y)/2 \in C$.

Theorem

Given a closed convex set C in a Hilbert space H, $\exists!$ point $x \in C$ that is closest to 0.

Proof.

We might as well assume $0 \notin C$. Choose a sequence $\{x_n\} \in C$ such that $\| x_n \| \to \inf_{x \in C} \| x \| := d$.

$$\frac{1}{2} \| x_n - x_m \|^2 = \| x_n \|^2 + \| x_m \|^2 - \frac{1}{2} \| x_n + x_m \|^2 \qquad \text{(parallelogram law)}$$

$$= \| x_n \|^2 + \| x_m \|^2 - \frac{1}{2}\left(4\left\|\frac{x_n + x_m}{2}\right\|^2\right) \qquad \text{(convexity)}$$

$$\leq \| x_n \|^2 + \| x_m \|^2 - 2d^2 \to 2d^2 - 2d^2 = 0 \quad \text{as } n, m \to \infty.$$

So, by taking limits as $n, m \to \infty$, $\lim_{n,m\to\infty} \| x_n - x_m \| = 0$, i.e., $\{x_n\}$ is Cauchy, so it converges to an $x_0 \in C$. Clearly:

$$\| x_0 \| = \| x_0 - x_n + x_n \| \leq \| x_0 - x_n \| + \| x_n \|,$$

so $\| x_0 \| = d$. If $\exists x_0 \neq y_0 \in C$ with $\| x_0 \| = \| y_0 \| = d$, then again by Parallelogram law:

$$\frac{\| x_0 - y_0 \|^2}{2} = -\frac{\| x_0 - y_0 \|^2}{2} + \| x_0 \|^2 + \| y_0 \|^2 \leq -2d^2 + 2d^2 = 0 \implies x_0 = y_0.$$

□

In practice, one needs to know the existence of a point in a given set $C \not\ni 0$ of a Hilbert space that is closest to 0, and if it exists, whether this point is unique.
One can get a satisfactory answer if C is a **closed convex** set in H. Closed in the sense that it contains all its limit points (i.e., $x_n \to x, x_n \in C \Longrightarrow x \in C$), and convex in the sense that $\forall x, y \in C$, the midpoint $(x+y)/2 \in C$.

Theorem

Given a closed convex set C in a Hilbert space H, $\exists!$ point $x \in C$ that is closest to 0.

Proof.

We might as well assume $0 \notin C$. Choose a sequence $\{x_n\} \in C$ such that $\| x_n \| \to \inf_{x\in C} \| x \| := d$.

$$\begin{aligned} \frac{1}{2} \| x_n - x_m \|^2 &= \| x_n \|^2 + \| x_m \|^2 - \frac{1}{2} \| x_n + x_m \|^2 && \text{(parallelogram law)} \\ &= \| x_n \|^2 + \| x_m \|^2 - \frac{1}{2}\left(4\left\|\frac{x_n + x_m}{2}\right\|^2\right) && \text{(convexity)} \\ &\leq \| x_n \|^2 + \| x_m \|^2 - 2d^2 \to 2d^2 - 2d^2 = 0 \quad \text{as } n, m \to \infty. \end{aligned}$$

So, by taking limits as $n, m \to \infty$, $\lim_{n,m\to\infty} \| x_n - x_m \| = 0$, i.e., $\{x_n\}$ is Cauchy, so it converges to an $x_0 \in C$. Clearly:

$$\| x_0 \| = \| x_0 - x_n + x_n \| \leq \| x_0 - x_n \| + \| x_n \|,$$

so $\| x_0 \| = d$. If $\exists x_0 \neq y_0 \in C$ with $\| x_0 \| = \| y_0 \| = d$, then again by Parallelogram law:

$$\frac{\| x_0 - y_0 \|^2}{2} = -\frac{\| x_0 - y_0 \|^2}{2} + \| x_0 \|^2 + \| y_0 \|^2 \leq -2d^2 + 2d^2 = 0 \quad \Longrightarrow x_0 = y_0.$$

□

In practice, one needs to know the existence of a point in a given set $C \not\ni 0$ of a Hilbert space that is closest to 0, and if it exists, whether this point is unique.

One can get a satisfactory answer if C is a **closed convex** set in H. Closed in the sense that it contains all its limit points (i.e., $x_n \to x, x_n \in C \Longrightarrow x \in C$), and convex in the sense that $\forall x, y \in C$, the midpoint $(x+y)/2 \in C$.

Theorem

Given a closed convex set C in a Hilbert space H, $\exists!$ point $x \in C$ that is closest to 0.

Proof.

We might as well assume $0 \notin C$. Choose a sequence $\{x_n\} \in C$ such that $\| x_n \| \to \inf_{x \in C} \| x \| := d$.

$$\begin{aligned}\frac{1}{2} \| x_n - x_m \|^2 &= \| x_n \|^2 + \| x_m \|^2 - \frac{1}{2} \| x_n + x_m \|^2 && \text{(parallelogram law)} \\ &= \| x_n \|^2 + \| x_m \|^2 - \frac{1}{2}\left(4\left\|\frac{x_n + x_m}{2}\right\|^2\right) && \text{(convexity)} \\ &\leq \| x_n \|^2 + \| x_m \|^2 - 2d^2 \to 2d^2 - 2d^2 = 0 \quad \text{as } n, m \to \infty.\end{aligned}$$

So, by taking limits as $n, m \to \infty$, $\lim_{n,m\to\infty} \| x_n - x_m \| = 0$, i.e., $\{x_n\}$ is Cauchy, so it converges to an $x_0 \in C$. Clearly:

$$\| x_0 \| = \| x_0 - x_n + x_n \| \leq \| x_0 - x_n \| + \| x_n \|,$$

so $\| x_0 \| = d$. If $\exists x_0 \neq y_0 \in C$ with $\| x_0 \| = \| y_0 \| = d$, then again by Parallelogram law:

$$\frac{\| x_0 - y_0 \|^2}{2} = -\frac{\| x_0 - y_0 \|^2}{2} + \| x_0 \|^2 + \| y_0 \|^2 \leq -2d^2 + 2d^2 = 0 \quad \Longrightarrow x_0 = y_0.$$

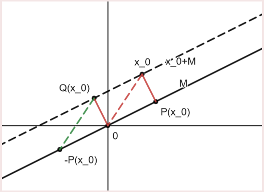
□

DISCUSSION

Now let us apply this result to problem of **finding the nearest point to a given point x_0 in a closed subspace M of H.**

By the theorem, the point we are seeking is the point x_0 minus the point with smallest norm of $x_0 + M$. Call this point $P(x_0) \in M$.

$$\| x_0 - P(x_0) \| = \inf_{m \in M} \| x_0 + m \| = \inf_{m \in M} \| x_0 - m \| .$$



Set $Q(x) := x - P(x)$. Let us look at the properties of P:

For $m \in M$ consider:

$$\| Q(x_0) \|^2 \leq \| Q(x_0) + \lambda m \|^2 = \| Q(x_0) \|^2 + 2\Re e \overline{\lambda} \langle Q(x_0), m \rangle + | \lambda |^2 \| m \|^2 .$$

If $\langle Q(x_0), m \rangle \neq 0$, choose $\lambda = t \frac{\langle Q(x_0), m \rangle}{|\langle Q(x_0), m \rangle|}$, $t \in \mathbb{R}$ to get:

$$0 \leq \frac{2t}{|t|} \, | \langle Q(x_0), m \rangle | + | t | \| m \|^2, \quad \forall t \in \mathbb{R} \setminus \{0\}.$$

Let $t \to 0^-$, since $\frac{t}{|t|} = -1$, we get $| \langle Q(x_0), m \rangle | = 0$. So $Q(x) \perp M \; \forall x \in H$. Hence:

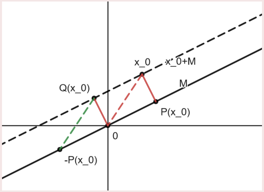
$$M \ni P(x + \lambda y) - \big(P(x) + \lambda P(y)\big) = P(x + \lambda y) - (x + \lambda y) - \big(P(x) - x\big) - \lambda\big(P(y) - y\big)$$
$$= -\big(Q(x + \lambda y) - Q(x) - \lambda Q(y)\big) \in M^{\perp}$$

DISCUSSION

Now let us apply this result to problem of **finding the nearest point to a given point x_0 in a closed subspace M of H.**

By the theorem, the point we are seeking is the point x_0 minus the point with smallest norm of $x_0 + M$. Call this point $P(x_0) \in M$.

$$\| x_0 - P(x_0) \| = \inf_{m \in M} \| x_0 + m \| = \inf_{m \in M} \| x_0 - m \| .$$



Set $Q(x) := x - P(x)$. Let us look at the properties of P:

For $m \in M$ consider:

$$\| Q(x_0) \|^2 \le \| Q(x_0) + \lambda m \|^2 = \| Q(x_0) \|^2 + 2\Re e \bar{\lambda} \langle Q(x_0), m \rangle + | \lambda |^2 \| m \|^2 .$$

If $\langle Q(x_0), m \rangle \neq 0$, choose $\lambda = t \frac{\langle Q(x_0), m \rangle}{|\langle Q(x_0), m \rangle|}$, $t \in \mathbb{R}$ to get:

$$0 \le \frac{2t}{|t|} \mid \langle Q(x_0), m \rangle \mid + \mid t \mid \| m \|^2, \quad \forall t \in \mathbb{R} \setminus \{0\}.$$

Let $t \to 0^-$, since $\frac{t}{|t|} = -1$, we get $| \langle Q(x_0), m \rangle | = 0$. So $Q(x) \perp M \ \forall x \in H$. Hence:

$$\begin{aligned} M \ni P(x + \lambda y) - \big(P(x) + \lambda P(y)\big) &= P(x + \lambda y) - (x + \lambda y) - \big(P(x) - x\big) - \lambda\big(P(y) - y\big) \\ &= -\big(Q(x + \lambda y) - Q(x) - \lambda Q(y)\big) \in M^\perp \end{aligned}$$

DISCUSSION (Cont.)

By the very definition

$$P^2(x) = P(P(x)) = P(x).$$

$$\| Q(P(x)) \| = \inf_{m \in M} \| P(x) - m \| = 0 \quad \Rightarrow QP(x) = 0.$$

Similarly,

$$\begin{aligned} & Q(x) = P(Q(x)) + Q(Q(x)) \\ \Longrightarrow \quad & M \ni P(Q(x)) = Q(x) - Q^2(x) \in M^{\perp} \\ \Longrightarrow \quad & Q^2 = Q, PQ = 0. \end{aligned}$$

In particular:

$$\begin{aligned} \langle P(x), y \rangle &= \langle P(x), P(y) \rangle \\ \langle x, P(y) \rangle &= \langle P(x), P(y) \rangle \end{aligned} \quad \text{and} \quad \begin{aligned} \langle P(x), y \rangle &= \langle x, P(y) \rangle \\ \langle Q(x), y \rangle &= \langle x, Q(y) \rangle \end{aligned}$$

For $x \in H$, $x = P(x) + Q(x)$.

Then:

$$\langle x, x \rangle = \langle P(x), P(x) \rangle + \langle Q(x), Q(x) \rangle \quad \Rightarrow \| P(x) \| \leq \| x \|, \| Q(x) \| \leq \| x \| .$$

We will summarize our findings in:

DISCUSSION (Cont.)

By the very definition

$$P^2(x) = P(P(x)) = P(x).$$

$$\| Q(P(x)) \| = \inf_{m \in M} \| P(x) - m \| = 0 \quad \Rightarrow QP(x) = 0.$$

Similarly,

$$\begin{aligned} & Q(x) = P(Q(x)) + Q(Q(x)) \\ \Longrightarrow \quad & M \ni P(Q(x)) = Q(x) - Q^2(x) \in M^{\perp} \\ \Longrightarrow \quad & Q^2 = Q, PQ = 0. \end{aligned}$$

In particular:

$$\begin{aligned} \langle P(x), y \rangle &= \langle P(x), P(y) \rangle \\ \langle x, P(y) \rangle &= \langle P(x), P(y) \rangle \end{aligned} \quad \text{and} \quad \begin{aligned} \langle P(x), y \rangle &= \langle x, P(y) \rangle \\ \langle Q(x), y \rangle &= \langle x, Q(y) \rangle \end{aligned}$$

For $x \in H$, $x = P(x) + Q(x)$.

Then:

$$\langle x, x \rangle = \langle P(x), P(x) \rangle + \langle Q(x), Q(x) \rangle \quad \Rightarrow \| P(x) \| \leq \| x \|, \| Q(x) \| \leq \| x \| .$$

We will summarize our findings in:

Theorem (Main Theorem)

Let $(H, \langle\,,\rangle)$ be a Hilbert space and M a closed subset of H. Then:

- *Every element x of H decomposes uniquely as $x = x_M + x_{M^\perp}$, where $x_M \in M$ and $x_{M^\perp} \in M^\perp$. In other words $H = M \oplus M^\perp$.*
- *There exist continuous linear operators $P : H \to M$, $Q : H \to M^\perp$ with*

$$\| P(x) \|^2 + \| Q(x) \|^2 = \| x \|^2$$

  such that $\forall x \in H$,

$$x = P(x) + Q(x)$$

  is the unique decomposition of x into M and $M^\perp$ above, i.e., $P + Q = I =$Identity.
- *$P^2 = P$, $PQ = 0$, $Q^2 = Q$ and*

$$\langle P(x), y\rangle = \langle x, P(y)\rangle, \quad \langle Q(x), y\rangle = \langle x, Q(y)\rangle$$

- *The operator Q satisfies the quantitative expression:*

$$\| Q(x) \| = \inf_{m \in M} \| x - m \| = \textit{distance of } x \textit{ to } M.$$

Terminology

We will call the operators P and Q above as **projections** onto M and $M^\perp$, respectively.

Theorem (Main Theorem)

Let $(H, \langle\, ,\rangle)$ be a Hilbert space and M a closed subset of H. Then:

- *Every element x of H decomposes uniquely as $x = x_M + x_{M^\perp}$, where $x_M \in M$ and $x_{M^\perp} \in M^\perp$. In other words $H = M \oplus M^\perp$.*
- *There exist continuous linear operators $P : H \to M$, $Q : H \to M^\perp$ with*

$$\| P(x) \|^2 + \| Q(x) \|^2 = \| x \|^2$$

such that $\forall x \in H$,

$$x = P(x) + Q(x)$$

is the unique decomposition of x into M and $M^\perp$ above, i.e., $P + Q = I =$Identity.

- *$P^2 = P$, $PQ = 0$, $Q^2 = Q$ and*

$$\langle P(x), y\rangle = \langle x, P(y)\rangle, \quad \langle Q(x), y\rangle = \langle x, Q(y)\rangle$$

- *The operator Q satisfies the quantitative expression:*

$$\| Q(x) \| = \inf_{m \in M} \| x - m \| = \textit{distance of } x \textit{ to } M.$$

Terminology

We will call the operators P and Q above as **projections** onto M and $M^\perp$, respectively.

Theorem (Riesz Representation Theorem)

Let $f : H \to \mathbb{C}$ be a continuous linear operator. Then there exists a unique element x_f of H with $f(x) = \langle x, x_f \rangle$.
Conversely, for any $y \in H$, the assignment $x \to \langle x, y \rangle$ defines a continuous linear operator from H into $\mathbb{C}$.

Proof.

Consider a continuous linear operator $f : H \to \mathbb{C}$. The kernel K of f, i.e., $K = \{x : f(x) = 0\}$ is a closed subspace of H since f is continuous. Consider a nonzero element $x_0 \in K^\perp$ (if no such element exists, we can take $x_f = 0$). For any $x \in H$, the decomposition

$$x = \left(x - \frac{f(x)}{f(x_0)} x_0\right) + \left(\frac{f(x)}{f(x_0)} x_0\right)$$

is the unique decomposition of x into $K^\perp$ and K since $x - \frac{f(x)}{f(x_0)} x_0 \in K$. So:

$$\left\langle x, \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0 \right\rangle = \left\langle \frac{f(x)}{f(x_0)} x_0, \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0 \right\rangle = f(x) \frac{f(x_0)}{f(x_0)} \frac{\langle x_0, x_0 \rangle}{\|x_0\|^2} = f(x).$$

It follows $f(x) = \langle x, x_f \rangle$ with $x_f = \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0$.
If there exists another $x_{\tilde{f}}$ with $f(x) = \langle x, x_{\tilde{f}} \rangle$, then $\langle x, x_f - x_{\tilde{f}} \rangle = 0 \ \forall x \in H$. In particular, $\langle x_f - x_{\tilde{f}}, x_f - x_{\tilde{f}} \rangle = 0$, which implies $x_f = x_{\tilde{f}}$.
The converse part of the theorem follows readily from the inequality

$$| \langle x, y \rangle | \leq \| x \| \| y \| \qquad \forall x, y \in H.$$

□

Theorem (Riesz Representation Theorem)

Let $f : H \to \mathbb{C}$ be a continuous linear operator. Then there exists a unique element x_f of H with $f(x) = \langle x, x_f \rangle$.
Conversely, for any $y \in H$, the assignment $x \to \langle x, y \rangle$ defines a continuous linear operator from H into $\mathbb{C}$.

Proof.

Consider a continuous linear operator $f : H \to \mathbb{C}$. The kernel K of f, i.e., $K = \{x : f(x) = 0\}$ is a closed subspace of H since f is continuous. Consider a nonzero element $x_0 \in K^\perp$ (if no such element exists, we can take $x_f = 0$). For any $x \in H$, the decomposition

$$x = \left(x - \frac{f(x)}{f(x_0)} x_0\right) + \left(\frac{f(x)}{f(x_0)} x_0\right)$$

is the unique decomposition of x into $K^\perp$ and K since $x - \frac{f(x)}{f(x_0)} x_0 \in K$. So:

$$\left\langle x, \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0 \right\rangle = \left\langle \frac{f(x)}{f(x_0)} x_0, \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0 \right\rangle = f(x) \frac{f(x_0)}{f(x_0)} \frac{\langle x_0, x_0 \rangle}{\|x_0\|^2} = f(x).$$

It follows $f(x) = \langle x, x_f \rangle$ with $x_f = \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0$.
If there exists another $x_{\tilde{f}}$ with $f(x) = \langle x, x_{\tilde{f}} \rangle$, then $\langle x, x_f - x_{\tilde{f}} \rangle = 0 \ \forall x \in H$. In particular, $\langle x_f - x_{\tilde{f}}, x_f - x_{\tilde{f}} \rangle = 0$, which implies $x_f = x_{\tilde{f}}$.
The converse part of the theorem follows readily from the inequality

$$| \langle x, y \rangle | \leq \| x \| \| y \| \quad \forall x, y \in H.$$

□

DISCUSSION

Now, given a continuous linear operator $T : H_1 \to H_2$, define an operator $T^* : H_2 \to H_1$ as follows: For any $y \in H_2$, T^*y is the unique element of H_1 such that for all $x \in H_1$

$$\langle Tx, y \rangle = \langle x, T^* y \rangle. \tag{3}$$

Such a T^*y exists because the assignment $x \to \langle Tx, y \rangle$ is a continuous linear operator from H_1 into $\mathbb{C}$ (since $| \langle Tx, y \rangle | \leq \| Tx \| \| y \| \leq c \| x \| \| y \|$ for some $c > 0$). So in view of the Riesz Representation Theorem there exists unique T^*y such that the Equation (3) above holds. This assignment is certainly linear and continuous:

$$\begin{aligned} \langle x, T^*(y_1 + \lambda y_2) \rangle &= \langle Tx, y_1 + \lambda y_2 \rangle = \langle Tx, y_1 \rangle + \overline{\lambda} \langle Tx, y_2 \rangle \\ &= \langle x, T^* y_1 \rangle + \overline{\lambda} \langle x, T^* y_2 \rangle = \langle x, T^* y_1 + \lambda T^* y_2 \rangle \quad \forall x; \end{aligned}$$

$$\begin{aligned} \| T^* y \|^2 &= | \langle T^* y, T^* y \rangle | = | \langle TT^* y, y \rangle | \leq c \| T^* y \| \| y \|, \\ \| T^* y \| &\leq c \| y \| . \end{aligned}$$

Definition

$T^* : H_2 \to H_1$ is a continuous linear operator and is called the **adjoint** of T.

If $T = T^*$ (defined on $H = H_1 = H_2$), then the operator is called **self-adjoint**. (Note that the projections in the Main Theorem are self adjoint.)

An operator $U : H \to H$ is called **unitary** in case $UU^* = I$. A unitary operator satisfies $\langle Ux, Uy \rangle = \langle x, y \rangle$, i.e., U preserves the inner product of elements.

DISCUSSION

Now, given a continuous linear operator $T : H_1 \to H_2$, define an operator $T^* : H_2 \to H_1$ as follows: For any $y \in H_2$, $T^* y$ is the unique element of H_1 such that for all $x \in H_1$

$$\boxed{\langle Tx, y\rangle = \langle x, T^* y\rangle.} \tag{3}$$

Such a $T^* y$ exists because the assignment $x \to \langle Tx, y\rangle$ is a continuous linear operator from H_1 into $\mathbb{C}$ (since $| \langle Tx, y\rangle | \leq \| Tx \| \| y \| \leq c \| x \| \| y \|$ for some $c > 0$). So in view of the Riesz Representation Theorem there exists unique $T^* y$ such that the Equation (3) above holds. This assignment is certainly linear and continuous:

$$\begin{aligned} \langle x, T^*(y_1 + \lambda y_2)\rangle &= \langle Tx, y_1 + \lambda y_2\rangle = \langle Tx, y_1\rangle + \overline{\lambda}\langle Tx, y_2\rangle \\ &= \langle x, T^* y_1\rangle + \overline{\lambda}\langle x, T^* y_2\rangle = \langle x, T^* y_1 + \lambda T^* y_2\rangle \quad \forall x; \end{aligned}$$

$$\begin{aligned} \| T^* y \|^2 &= | \langle T^* y, T^* y\rangle | = | \langle TT^* y, y\rangle | \leq c \| T^* y \| \| y \|, \\ \| T^* y \| &\leq c \| y \| . \end{aligned}$$

Definition

$T^* : H_2 \to H_1$ is a continuous linear operator and is called the **adjoint** of T.

If $T = T^*$ (defined on $H = H_1 = H_2$), then the operator is called **self-adjoint**. (Note that the projections in the Main Theorem are self adjoint.)

An operator $U : H \to H$ is called **unitary** in case $UU^* = I$. A unitary operator satisfies $\langle Ux, Uy\rangle = \langle x, y\rangle$, i.e., U preserves the inner product of elements.

Theorem

(0) *For a subspace $M \subseteq H$, $(M^{\perp})^{\perp} = \overline{M}$, where $\overline{M}$ is the closure of M.*

For a continuous linear operator $T : H_1 \to H_2$:

(1) $Ker(T) = Range(T^*)^{\perp}$

(2) $\overline{Range(T^*)} = Ker(T)^{\perp}$

(3) $T^{**} = T$

(4) $Ker(T^*) = Range(T)^{\perp} = \overline{Range(T)}^{\perp}$.

Proof.

(0): Recall that $M^{\perp} = \{x \in H : \langle x, m\rangle = 0\ \forall m \in M\}$.
Clearly $M \subseteq (M^{\perp})^{\perp}$ and since $(M^{\perp})^{\perp}$ is closed $\overline{M} \subseteq (M^{\perp})^{\perp}$.
For the other side, first note that $M^{\perp} = (\overline{M})^{\perp}$ since

$$\forall b \in \overline{M},\ \exists x_n \in M, n = 1, 2, \cdots \text{ such that } x_n \to b.$$

So for $a \in M^{\perp}$, $\langle a, x_n\rangle \to \langle a, b\rangle$ by continuity. Hence $\langle a, b\rangle = 0$.
Let $\alpha \in (M^{\perp})^{\perp} = (\overline{M})^{\perp\perp}$. By the Main Theorem $\alpha = \alpha_1 + \alpha_2$, with $\alpha_1 \in \overline{M}$, $\alpha_2 \in (\overline{M})^{\perp}$. So:

$$0 = \langle \alpha, \alpha_2\rangle = \langle \alpha_1, \alpha_2\rangle + \| \alpha_2 \|^2 = \| \alpha_2 \|^2 \quad \Rightarrow \alpha = \alpha_1 \in \overline{M}.$$

(1): $\xi \in Ker(T)$: $\langle \xi, T^* a\rangle = \langle T\xi, a\rangle = 0 \Rightarrow Ker(T) \subseteq R(T^*)^{\perp}$.
$\xi \in R(T^*)^{\perp} : \langle T\xi, \eta\rangle = \langle \xi, T^*\eta\rangle = 0\ \forall n \quad \therefore\ \xi \in Ker(T)$.

(3): $\langle T^{**}x, y\rangle = \overline{\langle y, (T^*)^* x\rangle} = \overline{\langle T^* y, x\rangle} = \overline{\langle y, Tx\rangle} = \langle Tx, y\rangle \Rightarrow\ T^{**}x = Tx\ \forall x$. □

Theorem

(0) *For a subspace $M \subseteq H$, $(M^{\perp})^{\perp} = \overline{M}$, where $\overline{M}$ is the closure of M.*

For a continuous linear operator $T : H_1 \to H_2$:

(1) $Ker(T) = Range(T^*)^{\perp}$

(2) $\overline{Range(T^*)} = Ker(T)^{\perp}$

(3) $T^{**} = T$

(4) $Ker(T^*) = Range(T)^{\perp} = \overline{Range(T)}^{\perp}$.

Proof.

(0): Recall that $M^{\perp} = \{x \in H : \langle x, m \rangle = 0 \, \forall m \in M\}$.
Clearly $M \subseteq (M^{\perp})^{\perp}$ and since $(M^{\perp})^{\perp}$ is closed $\overline{M} \subseteq (M^{\perp})^{\perp}$.
For the other side, first note that $M^{\perp} = (\overline{M})^{\perp}$ since

$$\forall b \in \overline{M}, \, \exists x_n \in M, n = 1, 2, \cdots \text{ such that } x_n \to b.$$

So for $a \in M^{\perp}$, $\langle a, x_n \rangle \to \langle a, b \rangle$ by continuity. Hence $\langle a, b \rangle = 0$.
Let $\alpha \in (M^{\perp})^{\perp} = (\overline{M})^{\perp\perp}$. By the Main Theorem $\alpha = \alpha_1 + \alpha_2$, with $\alpha_1 \in \overline{M}$, $\alpha_2 \in (\overline{M})^{\perp}$. So:

$$0 = \langle \alpha, \alpha_2 \rangle = \langle \alpha_1, \alpha_2 \rangle + \| \alpha_2 \|^2 = \| \alpha_2 \|^2 \quad \Rightarrow \alpha = \alpha_1 \in \overline{M}.$$

(1): $\xi \in Ker(T)$: $\langle \xi, T^* a \rangle = \langle T\xi, a \rangle = 0 \Rightarrow Ker(T) \subseteq R(T^*)^{\perp}$.
$\xi \in R(T^*)^{\perp} : \langle T\xi, \eta \rangle = \langle \xi, T^* \eta \rangle = 0 \, \forall n \quad \therefore \, \xi \in Ker(T)$.

(3): $\langle T^{**} x, y \rangle = \overline{\langle y, (T^*)^* x \rangle} = \overline{\langle T^* y, x \rangle} = \overline{\langle y, Tx \rangle} = \langle Tx, y \rangle \Rightarrow \; T^{**} x = Tx \, \forall x$. □

We will look at the equation $Tu = f$ for $T : H_1 \to H_2$ continuous linear operator and f, a given element of H_2.

Given T, H_1, H_2 and $f \in H_2$ as above, consider the condition:

$$\exists C > 0 : \quad | \langle f, x \rangle | \leq C \parallel T^* x \parallel \quad \forall x \in H_2 \tag{**}$$

If the equation $Tu = f$ has a solution, then for any $x \in H_2$:

$$| \langle f, x \rangle | = | \langle Tu, x \rangle | = | \langle u, T^* x \rangle | \leq \parallel u \parallel \parallel T^* x \parallel .$$

So (**) is satisfied with $C \geq \parallel u \parallel$.

On the other hand, if (**) is satisfied then on $R(T^*)$ define an operator S via $S(T^* v) := \langle v, f \rangle$. This assignment is well-defined since if $T^* v_1 = T^* v_2$ then (**) implies

$$| \langle f, v_1 \rangle - \langle f, v_2 \rangle | \leq C \parallel T^* v_1 - T^* v_2 \parallel = 0 \quad \Rightarrow \quad \langle v_1, f \rangle = \langle v_2, f \rangle .$$

Moreover by (**) we have $| S(T^* v) | \leq C \parallel T^* v \parallel$, so S is continuous and linear on $R(T^*)$.

As a general rule, S extends to $\overline{R(T^*)}$ by defining the extension for a given x, by

$$S(x) := \lim_{n \to \infty} S(x_n)$$

for some sequence $\{x_n\} \in R(T^*)$ that converges to x.

We will look at the equation $Tu = f$ for $T : H_1 \to H_2$ continuous linear operator and f, a given element of H_2.

Given T, H_1, H_2 and $f \in H_2$ as above, consider the condition:

$$\exists C > 0 : \quad |\langle f, x\rangle| \leq C \| T^* x \| \quad \forall x \in H_2 \tag{**}$$

If the equation $Tu = f$ has a solution, then for any $x \in H_2$:

$$|\langle f, x\rangle| = |\langle Tu, x\rangle| = |\langle u, T^* x\rangle| \leq \| u \| \| T^* x \| .$$

So (**) is satisfied with $C \geq \| u \|$.

On the other hand, if (**) is satisfied then on $R(T^*)$ define an operator S via $S(T^* v) := \langle v, f\rangle$. This assignment is well-defined since if $T^* v_1 = T^* v_2$ then (**) implies

$$|\langle f, v_1\rangle - \langle f, v_2\rangle| \leq C \| T^* v_1 - T^* v_2 \| = 0 \quad \Rightarrow \quad \langle v_1, f\rangle = \langle v_2, f\rangle.$$

Moreover by (**) we have $|S(T^* v)| \leq C \| T^* v \|$, so S is continuous and linear on $R(T^*)$.

As a general rule, S extends to $\overline{R(T^*)}$ by defining the extension for a given x, by

$$S(x) := \lim_{n\to\infty} S(x_n)$$

for some sequence $\{x_n\} \in R(T^*)$ that converges to x.

To see that this procedure does not depend upon the sequence chosen, suppose $x'_n \to x$ is another sequence in $R(T^*)$. Then since

$$\| S(x_n - x'_n) \| \leq C \| x_n - x'_n \|,$$

$\{S(x_n)\}$ and $\{S(x'_n)\}$ converges to the same point in $\mathbb{C}$, and since

$$| S(x_n) - S(x_m) | \leq C \| x_n - x_m \| \quad \forall n, m,$$

plainly $S(x_n)$ converges.

Moreover $\forall n$,

$$| S(x_n) | \leq C \| x_n \| \quad \Rightarrow \quad \| Sx \| \leq C \| x \| .$$

That is, S is continuous on $\overline{R(T^*)}$.

Let P be the projection on $\overline{R(T^*)}$ and consider $S \circ P(x)$. This is a continuous linear function from H_1 into $\mathbb{C}$ with

$$| S \circ P(x) | \leq C \| P(x) \| \leq C \| x \| .$$

So by Riesz Representation Theorem $\exists u \in H_2$ that satisfies

$$S \circ P(x) = \langle x, u \rangle \quad \forall x \in H_1 .$$

In particular for $x = T^* v$ we have

$$\langle v, f \rangle = \langle T^* v, u \rangle = \langle v, Tu \rangle, \qquad \text{or}$$
$$\langle f, v \rangle = \langle Tu, v \rangle, \forall v \quad \Rightarrow \quad f = Tu$$

Moreover we have an estimate on u, namely $\| u \| \leq C$.

To see that this procedure does not depend upon the sequence chosen, suppose $x'_n \to x$ is another sequence in $R(T^*)$. Then since

$$\| S(x_n - x'_n) \| \leq C \| x_n - x'_n \|,$$

$\{S(x_n)\}$ and $\{S(x'_n)\}$ converges to the same point in $\mathbb{C}$, and since

$$| S(x_n) - S(x_m) | \leq C \| x_n - x_m \| \quad \forall n, m,$$

plainly $S(x_n)$ converges.

Moreover $\forall n$,

$$| S(x_n) | \leq C \| x_n \| \quad \Rightarrow \quad \| Sx \| \leq C \| x \| .$$

That is, S is continuous on $\overline{R(T^*)}$.

Let P be the projection on $\overline{R(T^*)}$ and consider $S \circ P(x)$. This is a continuous linear function from H_1 into $\mathbb{C}$ with

$$| S \circ P(x) | \leq C \| P(x) \| \leq C \| x \| .$$

So by Riesz Representation Theorem $\exists u \in H_2$ that satisfies

$$S \circ P(x) = \langle x, u \rangle \quad \forall x \in H_1 .$$

In particular for $x = T^* v$ we have

$$\langle v, f \rangle = \langle T^* v, u \rangle = \langle v, Tu \rangle, \qquad \text{or}$$
$$\langle f, v \rangle = \langle Tu, v \rangle, \forall v \quad \Rightarrow \quad f = Tu$$

Moreover we have an estimate on u, namely $\| u \| \leq C$.

To summarize:

For a given $f \in H_2$ the equation $T(u) = f$ has a solution if and only if (**) holds.

In practice, sometimes explicit knowledge of T^* allows one to get a stronger estimate:

$$\exists C > 0 : \quad \| x \| \leq C \| T^* x \| \quad \forall x \in R(T). \tag{***}$$

This yields for $f \in \overline{R(T)}$ and $x \in \overline{R(T)}$:

$$| \langle f, x \rangle | \leq \| f \| \| x \| \leq C \| f \| \| T^* x \|,$$

which implies

$$\forall f \in \overline{R(T)} \quad \exists u \text{ with } Tu = f \text{ and } \| u \| \leq C \| f \| .$$

In particular, this implies that $R(T)$ is closed.

To summarize:

For a given $f \in H_2$ the equation $T(u) = f$ has a solution if and only if (**) holds.

In practice, sometimes explicit knowledge of T^* allows one to get a stronger estimate:

$$\exists C > 0 : \quad \| x \| \leq C \| T^* x \| \quad \forall x \in R(T). \tag{***}$$

This yields for $f \in \overline{R(T)}$ and $x \in \overline{R(T)}$:

$$| \langle f, x \rangle | \leq \| f \| \| x \| \leq C \| f \| \| T^* x \|,$$

which implies

$$\forall f \in \overline{R(T)} \quad \exists u \text{ with } Tu = f \text{ and } \| u \| \leq C \| f \| .$$

In particular, this implies that $R(T)$ is closed.

We will close this part by generalizing a special feature of the space ℓ_2 to a large class of Hilbert spaces.

Recall that:

$$\ell_2 := \{(\zeta_n) \in \mathbb{C}^{\mathbb{N}} : \sum_{n=1}^{\infty} |\zeta_n|^2 < \infty\},$$

with inner product

$$\langle \zeta, \eta \rangle := \sum_{n=1}^{\infty} \zeta_n \overline{\eta_n},\ \zeta = (\zeta_n),\ \eta = (\eta_n).$$

Note that $\ell_2 = \overline{span\{e_n\}}$, the closure of all finite linear combinations of e_n's, where

$$e_n := (0, \cdots, 0, 1, 0, \cdots),$$

since for any $\zeta = (\zeta_n) \in \ell_2$,

$$\| \zeta - \sum_{n=1}^{N} \zeta_n e_n \| = \sum_{n=N+1}^{\infty} |\zeta_n|^2 \to 0 \text{ as } N \to \infty.$$

We will call a Hilbert space H **separable** in case there exists a countable set of elements such that the closure of the span of these elements is H.

We will close this part by generalizing a special feature of the space ℓ_2 to a large class of Hilbert spaces.

Recall that:

$$\ell_2 := \{(\zeta_n) \in \mathbb{C}^{\mathbb{N}} : \sum_{n=1}^{\infty} | \zeta_n |^2 < \infty\},$$

with inner product

$$\langle \zeta, \eta \rangle := \sum_{n=1}^{\infty} \zeta_n \overline{\eta_n},\ \zeta = (\zeta_n),\ \eta = (\eta_n).$$

Note that $\ell_2 = \overline{span\{e_n\}}$, the closure of all finite linear combinations of e_n's, where

$$e_n := (0, \cdots, 0, 1, 0, \cdots),$$

since for any $\zeta = (\zeta_n) \in \ell_2$,

$$\| \zeta - \sum_{n=1}^{N} \zeta_n e_n \| = \sum_{n=N+1}^{\infty} | \zeta_n |^2 \to 0 \text{ as } N \to \infty.$$

We will call a Hilbert space H **separable** in case there exists a countable set of elements such that the closure of the span of these elements is H.

Note that in ℓ_2 every element ζ can be expressed as:

$$\zeta = \sum_{n=1}^{\infty} \langle \zeta_n, e_n \rangle e_n,$$

the series converging in ℓ_2.

Now, if H is a separable Hilbert space with $\overline{span\{x_n\}} = H$, by Gram-Schmidt algorithm we can get another sequence (y_n) with $\overline{span\{y_n\}} = H$ and $\langle y_n, y_m \rangle = \delta_{n,m}$.

For a given $x \in H$ form:

$$x_N := \sum_{n=1}^{N} \langle x, y_n \rangle y_n, \quad N \in \mathbb{N}.$$

Since $\langle x - x_N, y_i \rangle = 0$ for $i = 1, \cdots, N$;

$$x - x_N \in span\{y_1, \cdots, y_N\}^{\perp} = \overline{span\{y_1, \cdots, y_N\}}^{\perp}.$$

Then $x = (x - x_N) + x_N$ is the unique decomposition of x given by the Main Theorem.

In particular:

$$\langle x - x_N + x_N, x - x_N + x_N \rangle = \| x \|^2 = \| x - x_N \|^2 + \| x_N \|^2 .$$

Note that in ℓ_2 every element ζ can be expressed as:

$$\zeta = \sum_{n=1}^{\infty} \langle \zeta_n, e_n \rangle e_n,$$

the series converging in ℓ_2.

Now, if H is a separable Hilbert space with $\overline{span\{x_n\}} = H$, by Gram-Schmidt algorithm we can get another sequence (y_n) with $\overline{span\{y_n\}} = H$ and $\langle y_n, y_m \rangle = \delta_{n,m}$.

For a given $x \in H$ form:

$$x_N := \sum_{n=1}^{N} \langle x, y_n \rangle y_n, \quad N \in \mathbb{N}.$$

Since $\langle x - x_N, y_i \rangle = 0$ for $i = 1, \cdots, N$;

$$x - x_N \in span\{y_1, \cdots, y_N\}^{\perp} = \overline{span\{y_1, \cdots, y_N\}}^{\perp}.$$

Then $x = (x - x_N) + x_N$ is the unique decomposition of x given by the Main Theorem.

In particular:

$$\langle x - x_N + x_N, x - x_N + x_N \rangle = \| x \|^2 = \| x - x_N \|^2 + \| x_N \|^2 .$$

So we can draw two conclusions from this and the Main Theorem:

1) $\| x - x_N \|^2 =$ distance of x to $\overline{span\{x_1, \cdots, x_N\}} \to 0$ as $N \to \infty$ by our assumption,
2) $\| x_N \|^2 = \sum_{n=1}^{N} | \langle x, y_n \rangle |^2 \leq \| x \|^2$, so $\{\langle x, y_n \rangle\} \in \ell_2$.

It follows that every $x \in H$ can be expanded in H as:

$$x = \sum_{n=1}^{\infty} \langle x, y_n \rangle y_n \quad \text{with} \quad \{\langle x, y_n \rangle\} \in \ell_2.$$

Conversely, if $\{\lambda_n\}$ is in ℓ_2, the series $\sum_{n=1}^{\infty} \lambda_n y_n$ converges in H, since (if $N < M$)

$$\left\| \sum_{n=1}^{N} \lambda_n y_n - \sum_{n=1}^{M} \lambda_n y_n \right\|^2 = \left\| \sum_{n=N+1}^{M} \lambda_n y_n \right\|^2 = \sum_{n=N+1}^{M} | \lambda_n |^2 \to 0 \text{ as } N, M \to \infty.$$

So $\sum_{n=1}^{N} \lambda_n y_n$ converges to a point ζ in H. In particular, for each y_s,

$$\left\langle \sum_{n=1}^{N} \lambda_n y_n, y_s \right\rangle = \lambda_s \text{ if } N \geq s, \text{ so } \langle \zeta, y_s \rangle = \lambda_s.$$

So $\zeta = \sum_{n=1}^{\infty} \lambda_n y_n$ and $\lambda_n = \langle \zeta, y_n \rangle$, $n = 1, \cdots$.

So we can draw two conclusions from this and the Main Theorem:

1) $\| x - x_N \|^2 =$ distance of x to $\overline{span\{x_1, \cdots, x_N\}} \to 0$ as $N \to \infty$ by our assumption,
2) $\| x_N \|^2 = \sum_{n=1}^{N} | \langle x, y_n \rangle |^2 \leq \| x \|^2$, so $\{\langle x, y_n \rangle\} \in \ell_2$.

It follows that every $x \in H$ can be expanded in H as:

$$x = \sum_{n=1}^{\infty} \langle x, y_n \rangle y_n \quad \text{with} \quad \{\langle x, y_n \rangle\} \in \ell_2.$$

Conversely, if $\{\lambda_n\}$ is in ℓ_2, the series $\sum_{n=1}^{\infty} \lambda_n y_n$ converges in H, since (if $N < M$)

$$\left\| \sum_{n=1}^{N} \lambda_n y_n - \sum_{n=1}^{M} \lambda_n y_n \right\|^2 = \left\| \sum_{n=N+1}^{M} \lambda_n y_n \right\|^2 = \sum_{n=N+1}^{M} | \lambda_n |^2 \to 0 \text{ as } N, M \to \infty.$$

So $\sum_{n=1}^{N} \lambda_n y_n$ converges to a point ζ in H. In particular, for each y_s,

$$\left\langle \sum_{n=1}^{N} \lambda_n y_n, y_s \right\rangle = \lambda_s \text{ if } N \geq s, \text{ so } \langle \zeta, y_s \rangle = \lambda_s.$$

So $\zeta = \sum_{n=1}^{\infty} \lambda_n y_n$ and $\lambda_n = \langle \zeta, y_n \rangle$, $n = 1, \cdots$.

To summarize

For a given separable H, $\exists\{y_n\}_{n=1}^{\infty}$ with $\langle y_n, y_m\rangle = \delta_{n,m}$ such that every $x \in H$ can be expanded uniquely in a series

$$x = \sum_{n=1}^{\infty} \lambda_n y_n, \quad \{\lambda_n\}_n \in \ell_2,\ \lambda_n = \langle x, y_n\rangle,\ n = 1, 2, \cdots .$$

Such a sequence $\{y_n\}$ will be referred as an **orthonormal basis**.

Continuing our discussion, it follows that there is an operator $T : H \to \ell_2$, $Tx := \{\langle x, y_n\rangle\}_n$ that is one to one and onto. Moreover, for $f, g \in H$:

$$\langle f, g\rangle = \lim_{N\to\infty}\left\langle \sum_{n=1}^{N}\langle f, y_n\rangle y_n, g\right\rangle = \lim_{N\to\infty}\sum_{n=1}^{N}\langle f, y_n\rangle\overline{\langle g, y_n\rangle} = \sum_{n=1}^{\infty}\langle f, y_n\rangle\overline{\langle g, y_n\rangle}.$$

The last equality is valid since if $N < M$

$$\left|\sum_{n=1}^{N}\langle f, y_n\rangle\overline{\langle g, y_n\rangle} - \sum_{n=1}^{M}\langle f, y_n\rangle\overline{\langle g, y_n\rangle}\right| \le \left|\sum_{n=N+1}^{M}\langle f, y_n\rangle\overline{\langle g, y_n\rangle}\right|$$

$$\le \left(\sum_{n=N+1}^{M} |\langle f, y_n\rangle|^{1/2}\right)^2 \left(\sum_{n=N+1}^{M} |\langle g, y_n\rangle|^2\right)^{1/2}$$

$$\to 0 \text{ as } N, M \to \infty.$$

To summarize

For a given separable H, $\exists\{y_n\}_{n=1}^{\infty}$ with $\langle y_n, y_m\rangle = \delta_{n,m}$ such that every $x \in H$ can be expanded uniquely in a series

$$x = \sum_{n=1}^{\infty} \lambda_n y_n, \quad \{\lambda_n\}_n \in \ell_2,\ \lambda_n = \langle x, y_n\rangle,\ n = 1, 2, \cdots .$$

Such a sequence $\{y_n\}$ will be referred as an **orthonormal basis**.

Continuing our discussion, it follows that there is an operator $T : H \to \ell_2$, $Tx := \{\langle x, y_n\rangle\}_n$ that is one to one and onto. Moreover, for $f, g \in H$:

$$\langle f, g\rangle = \lim_{N\to\infty}\left\langle \sum_{n=1}^{N} \langle f, y_n\rangle y_n, g\right\rangle = \lim_{N\to\infty} \sum_{n=1}^{N} \langle f, y_n\rangle\overline{\langle g, y_n\rangle} = \sum_{n=1}^{\infty} \langle f, y_n\rangle\overline{\langle g, y_n\rangle}.$$

The last equality is valid since if $N < M$

$$\begin{aligned}\left|\sum_{n=1}^{N} \langle f, y_n\rangle\overline{\langle g, y_n\rangle} - \sum_{n=1}^{M} \langle f, y_n\rangle\overline{\langle g, y_n\rangle}\right| &\leq \left|\sum_{n=N+1}^{M} \langle f, y_n\rangle\overline{\langle g, y_n\rangle}\right| \\ &\leq \left(\sum_{n=N+1}^{M} |\langle f, y_n\rangle|^{1/2}\right)^2 \left(\sum_{n=N+1}^{M} |\langle g, y_n\rangle|^2\right)^{1/2} \\ &\to 0 \text{ as } N, M \to \infty.\end{aligned}$$

Therefore

$$\| T(h) \| = \left(\sum_{n=1}^{\infty} | \langle h, y_n \rangle |^2 \right)^{1/2} = \| h \| .$$

More generally;

$$\langle T(h), T(g) \rangle = \langle f, g \rangle.$$

That is, T is unitary isomorphism from H onto ℓ_2.

The moral of the story is

In a separable Hilbert space H, one can introduce "coordinates" in ℓ_2 just like in $\mathbb{C}^n$ one introduces $x \leftrightarrow (x_1, \cdots, x_n)$, $x_n \in \mathbb{C}$ and work with these coordinates.

Therefore

$$\| T(h) \| = \left(\sum_{n=1}^{\infty} | \langle h, y_n \rangle |^2 \right)^{1/2} = \| h \| .$$

More generally;

$$\langle T(h), T(g) \rangle = \langle f, g \rangle.$$

That is, T is unitary isomorphism from H onto ℓ_2.

The moral of the story is

In a separable Hilbert space H, one can introduce "coordinates" in ℓ_2 just like in $\mathbb{C}^n$ one introduces $x \leftrightarrow (x_1, \cdots, x_n)$, $x_n \in \mathbb{C}$ and work with these coordinates.

Let $(X, \langle \, , \rangle)$ be an inner product space that is not complete. This inner product induces a metric on X as we have seen earlier via,

$$\| x - y \|^2 := \langle x - y, x - y \rangle.$$

Let us call a function $f : X \to \mathbb{C}$ *anti-linear* in case $f(x + \lambda y) = f(x) + \overline{\lambda} f(y)$ for every $x, y \in X$ and $\lambda \in \mathbb{C}$. Consider

$$\begin{aligned} X^* &= \{f : X \to \mathbb{C} : f \text{ is anti-linear and continuous}\} \\ &= \{f : X \to \mathbb{C} : f \text{ is anti-linear and } \exists C > 0 \text{ s.t. } \mid f(x) \mid \leq C \| x \| \}. \end{aligned}$$

Note that X^* is a subspace of the vector space of complex valued functions on X.

We can identify elements of X with a subset of X^* via $x \mapsto f_x$, $f_x(y) := \langle x, y \rangle$. Note that this is a one to one and linear assignment, i.e.,

$$f_{x_1} = f_{x_2} \implies \langle y, x_1 - x_2 \rangle = 0 \quad \forall y \quad \implies x_1 = x_2,$$

and

$$f_{x+ty} = f_x + t f_y \quad \text{for } x, y \in X, \; t \in \mathbb{C}.$$

Moreover

$$\sup_{\|y\| \leq 1} \mid f_x(y) \mid = \| x \| \quad \forall x \in X.$$

Choose a Cauchy sequence $\{x_n\}$ in X, then since

$$\mid f_{x_n}(t) - f_{x_m}(t) \mid = \mid \langle t, x_n - x_m \rangle \mid \leq \| x_n - x_m \| \| t \| \quad \forall t \in X,$$

$f_{x_n}(t)$ converges to a point in $\mathbb{C}$ as $n \to \infty$. Call this point $f(t)$.

The function $t \mapsto f(t)$ is clearly anti-linear since each f_{x_n} is, $n = 1, 2, \ldots$

Let $(X, \langle\, , \rangle)$ be an inner product space that is not complete. This inner product induces a metric on X as we have seen earlier via,

$$\| x - y \|^2 := \langle x - y, x - y \rangle.$$

Let us call a function $f : X \to \mathbb{C}$ *anti-linear* in case $f(x + \lambda y) = f(x) + \bar{\lambda} f(y)$ for every $x, y \in X$ and $\lambda \in \mathbb{C}$.

Consider

$$\begin{aligned} X^* &= \{f : X \to \mathbb{C} : f \text{ is anti-linear and continuous}\} \\ &= \{f : X \to \mathbb{C} : f \text{ is anti-linear and } \exists C > 0 \text{ s.t. } \mid f(x) \mid \leq C \| x \|\}. \end{aligned}$$

Note that X^* is a subspace of the vector space of complex valued functions on X.

We can identify elements of X with a subset of X^* via $x \mapsto f_x$, $f_x(y) := \langle x, y \rangle$. Note that this is a one to one and linear assignment, i.e.,

$$f_{x_1} = f_{x_2} \implies \langle y, x_1 - x_2 \rangle = 0 \quad \forall y \quad \implies x_1 = x_2,$$

and

$$f_{x+ty} = f_x + t f_y \quad \text{for } x, y \in X,\ t \in \mathbb{C}.$$

Moreover

$$\sup_{\|y\| \leq 1} \mid f_x(y) \mid = \| x \| \quad \forall x \in X.$$

Choose a Cauchy sequence $\{x_n\}$ in X, then since

$$\mid f_{x_n}(t) - f_{x_m}(t) \mid = \mid \langle t, x_n - x_m \rangle \mid \leq \| x_n - x_m \| \| t \| \quad \forall t \in X,$$

$f_{x_n}(t)$ converges to a point in $\mathbb{C}$ as $n \to \infty$. Call this point $f(t)$.

The function $t \mapsto f(t)$ is clearly anti-linear since each f_{x_n} is, $n = 1, 2, \cdots$.

Moreover;

$$| f(t) | \leq \lim_{n\to\infty} | f_{x_n}(t) | \leq \lim_{n\to\infty} | \langle x_n, t\rangle | \leq \| t \| \| x_n \| .$$

Since $\forall \varepsilon > 0 \exists N : \| x_n - x_m \| \leq \varepsilon$ if $n, m \geq N$,

$$\| x_n \| \leq \| x_n - x_N \| + \| x_N \| \leq \varepsilon + \| x_N \| = C.$$

So $f \in X^*$.

Note that a similar argument presented above shows that X^* with the metric

$$d(f,g) = \sup_{\|y\|\leq 1} | f(y) - g(y) |$$

is complete, that is, every Cauchy sequence in X^* converges.

Note that $\| x \| = d(0, f_x)$.

Now consider the closure of X in X^*. For an $f \in X^*$, define

$$\langle f, f_x\rangle := f(x), \quad x \in X.$$

Certainly, it is linear in f and anti-linear in f_x's, that is $\langle f, cf_x\rangle = \overline{c}\langle f, f_x\rangle$ for $c \in \mathbb{C}$.
For $f = f_y$, $\langle f_y, f_x\rangle = f_y(x) = \langle y, x\rangle$.

For a $g \in \overline{X}$, choose sequences $\{x_n\}$ and $\{y_n\}$ such that $f_{x_n} \to g$ and $f_{y_n} \to g$ in X^*.

Moreover;

$$| f(t) | \leq \lim_{n\to\infty} | f_{x_n}(t) | \leq \lim_{n\to\infty} | \langle x_n, t\rangle | \leq \| t \| \| x_n \| .$$

Since $\forall \varepsilon > 0 \exists N : \| x_n - x_m \| \leq \varepsilon$ if $n, m \geq N$,

$$\| x_n \| \leq \| x_n - x_N \| + \| x_N \| \leq \varepsilon + \| x_N \| = C.$$

So $f \in X^*$.

Note that a similar argument presented above shows that X^* with the metric

$$d(f, g) = \sup_{\|y\|\leq 1} | f(y) - g(y) |$$

is complete, that is, every Cauchy sequence in X^* converges.

Note that $\| x \| = d(0, f_x)$.

Now consider the closure of X in X^*. For an $f \in X^*$, define

$$\langle f, f_x\rangle := f(x), \quad x \in X.$$

Certainly, it is linear in f and anti-linear in f_x's, that is $\langle f, cf_x\rangle = \overline{c}\langle f, f_x\rangle$ for $c \in \mathbb{C}$.

For $f = f_y$, $\langle f_y, f_x\rangle = f_y(x) = \langle y, x\rangle$.

For a $g \in \overline{X}$, choose sequences $\{x_n\}$ and $\{y_n\}$ such that $f_{x_n} \to g$ and $f_{y_n} \to g$ in X^*.

Then for $f \in \overline{X}$ there exist a $C > 0$ such that,

$$\begin{aligned} | \langle f, f_{x_n} - f_{y_n} \rangle | &= | \langle f, f_{x_n} \rangle - \langle f, f_{y_n} \rangle | = | f(x_n) - f(y_n) | \\ &= | f(x_n - y_n) | \leq C \| x_n - y_n \| = Cd(f_{x_n}, f_{y_n}), \end{aligned}$$

and

$$\begin{aligned} | \langle f, f_{x_n} \rangle - \langle f, f_{x_m} \rangle | &= | \langle f, f_{x_n} - f_{x_m} \rangle | = | f(x_n - x_m) | \\ &\leq C \| x_n - x_m \| \leq Cd(f_{x_n}, f_{x_m}). \end{aligned}$$

Hence we can define

$$\langle f, g \rangle := \lim_{n\to\infty} \langle f, f_{x_n} \rangle, \quad \text{for } f_{x_n} \to g,$$

and this definition does not depend upon the sequence f_{x_n} chosen.

Clearly this assignment is sesquilinear.

Note that $\langle f, f \rangle = \lim_{n\to\infty} \langle f, f_{x_n} \rangle$ for a sequence $f_{x_n} \to f$ in X^*:

$$\begin{aligned} | f_{x_n}(x_n) - f(x_n) | &\leq \| x_n \| \, d(f_{x_n}, f) = d(0, f_{x_n}) d(f_{x_n}, f) \\ &\leq d(f_{x_n}, f)^2 + d(0, f) d(f_{x_n}, f). \end{aligned}$$

So $lim_{n\to\infty} | f_{x_n}(x_n) - f(x_n) | = 0$. Observe:

$$f(x_n) = -f_{x_n}(x_n) + f(x_n) + f_{x_n}(x_n) = f(x_n) - f_{x_n}(x_n) + \langle x_n, x_n \rangle.$$

So $\langle f, f \rangle = \lim_{n\to\infty} f(x_n) \geq 0$ and is zero if $\langle x_n, x_n \rangle \to 0$, which in view of

$$| f(y) | = \lim_{n\to\infty} | f_{x_n}(y) | = \lim_{n\to\infty} | \langle y, x_n \rangle | \leq \| y \| \| x_n \|,$$

implies $f \equiv 0$.

Then for $f \in \overline{X}$ there exist a $C > 0$ such that,

$$\begin{aligned} | \langle f, f_{x_n} - f_{y_n} \rangle | &=| \langle f, f_{x_n} \rangle - \langle f, f_{y_n} \rangle |=| f(x_n) - f(y_n) | \\ &=| f(x_n - y_n) |\leq C \parallel x_n - y_n \parallel= Cd(f_{x_n}, f_{y_n}), \end{aligned}$$

and

$$\begin{aligned} | \langle f, f_{x_n} \rangle - \langle f, f_{x_m} \rangle | &=| \langle f, f_{x_n} - f_{x_m} \rangle |=| f(x_n - x_m) | \\ &\leq C \parallel x_n - x_m \parallel\leq Cd(f_{x_n}, f_{x_m}). \end{aligned}$$

Hence we can define

$$\langle f, g \rangle := \lim_{n\to\infty} \langle f, f_{x_n} \rangle, \quad \text{for } f_{x_n} \to g,$$

and this definition does not depend upon the sequence f_{x_n} chosen.

Clearly this assignment is sesquilinear.

Note that $\langle f, f \rangle = \lim_{n\to\infty} \langle f, f_{x_n} \rangle$ for a sequence $f_{x_n} \to f$ in X^*:

$$\begin{aligned} | f_{x_n}(x_n) - f(x_n) | &\leq\parallel x_n \parallel d(f_{x_n}, f) = d(0, f_{x_n}) d(f_{x_n}, f) \\ &\leq d(f_{x_n}, f)^2 + d(0, f) d(f_{x_n}, f). \end{aligned}$$

So $lim_{n\to\infty} | f_{x_n}(x_n) - f(x_n) |= 0$. Observe:

$$f(x_n) = -f_{x_n}(x_n) + f(x_n) + f_{x_n}(x_n) = f(x_n) - f_{x_n}(x_n) + \langle x_n, x_n \rangle.$$

So $\langle f, f \rangle = \lim_{n\to\infty} f(x_n) \geq 0$ and is zero if $\langle x_n, x_n \rangle \to 0$, which in view of

$$| f(y) |= \lim_{n\to\infty} | f_{x_n}(y) |= \lim_{n\to\infty} | \langle y, x_n \rangle |\leq\parallel y \parallel \parallel x_n \parallel,$$

implies $f \equiv 0$.

It follows that $\langle \, , \rangle$ is an inner product on $\overline{X}$.

Moreover, if $\{f_n\}$ is a Cauchy sequence in $\overline{X}$ with respect to the topology coming from this inner product, since for $x \in X$:

$$\begin{aligned} | f_n(x) - f_m(x) | \leq | \langle f_n - f_m, f_x \rangle | \leq | \langle f_n - f_m, f_n - f_m \rangle |^{1/2} | \langle f_x, f_x \rangle |^{1/2} \\ \leq \langle f_n - f_m, f_n - f_m \rangle^{1/2} \langle x, x \rangle^{1/2}, \\ d(f_n, f_m) = \sup_{\|x\| \leq 1} | f_n - f_m | \leq \langle f_n - f_m, f_n - f_m \rangle^{1/2}, \end{aligned}$$

$\{f_n\}$ is a Cauchy sequence in $\overline{X}$ with respect to the original topology of $\overline{X}$.

Since $\overline{X}$ is complete $f_n \to f$ in this topology.

On the other hand, for a given $g \in \overline{X}$, choosing $f_{x_k} \to g$ in X^* we have:

$$\begin{aligned} | \langle g, g \rangle | = \lim_{k \to \infty} | \langle g, f_{x_k} \rangle | = \lim_{k \to \infty} | g(x_k) | \\ \leq \lim_{k \to \infty} d(0, g) \, \| x_k \| = \lim_{k \to \infty} d(0, g) \, d(0, f_{x_k}) = d(0, g)^2. \end{aligned}$$

So

$$| \langle f_n - f_m, f_n - f_m \rangle | \leq d(f_n, f)^2.$$

This implies that $f_n \to f$ in $(\overline{X}, \langle \, , \rangle)$.

It follows that $\langle\,,\rangle$ is an inner product on $\overline{X}$.

Moreover, if $\{f_n\}$ is a Cauchy sequence in $\overline{X}$ with respect to the topology coming from this inner product, since for $x \in X$:

$$\begin{aligned} | f_n(x) - f_m(x) | &\leq | \langle f_n - f_m, f_x \rangle | \leq | \langle f_n - f_m, f_n - f_m \rangle |^{1/2} | \langle f_x, f_x \rangle |^{1/2} \\ &\leq \langle f_n - f_m, f_n - f_m \rangle^{1/2} \langle x, x \rangle^{1/2}, \\ d(f_n, f_m) = \sup_{\|x\| \leq 1} | f_n - f_m | &\leq \langle f_n - f_m, f_n - f_m \rangle^{1/2}, \end{aligned}$$

$\{f_n\}$ is a Cauchy sequence in $\overline{X}$ with respect to the original topology of $\overline{X}$.

Since $\overline{X}$ is complete $f_n \to f$ in this topology.

On the other hand, for a given $g \in \overline{X}$, choosing $f_{x_k} \to g$ in X^* we have:

$$\begin{aligned} | \langle g, g \rangle | &= \lim_{k \to \infty} | \langle g, f_{x_k} \rangle | = \lim_{k \to \infty} | g(x_k) | \\ &\leq \lim_{k \to \infty} d(0, g) \, \| x_k \| = \lim_{k \to \infty} d(0, g) \, d(0, f_{x_k}) = d(0, g)^2. \end{aligned}$$

So

$$| \langle f_n - f_m, f_n - f_m \rangle | \leq d(f_n, f)^2.$$

This implies that $f_n \to f$ in $(\overline{X}, \langle\,,\rangle)$.

So $(\overline{X}, \langle\, , \rangle)$ is a Hilbert space, it contains $(X, \langle\, , \rangle)$, $\langle\, , \rangle$ induces the inner product on X, moreover, the closure of X is the full space $\overline{X}$.

If $(H, \langle\langle\, , \rangle\rangle)$ is another Hilbert space enjoining the above mentioned properties of $(\overline{X}, \langle\, , \rangle)$, then the identity operator on $(X, \langle\, , \rangle)$ plainly extends to a 1-1, onto unitary operator from $\overline{X}$ to H.

The unique Hilbert space satisfying the above mentioned properties is called the *completion* of $(X, \langle\, , \rangle)$.

Now, going back to our construction, if $(X, \langle\, , \rangle)$ is a vector space of functions on a set T where point evaluations are continuous, then for f_x, set $f_x(t) := x(t)$ and if $f_{x_n} \to f$ in $(\overline{X}, \langle\, , \rangle)$, then we propose to set $f(t) = \lim_{n\to\infty} x_n(t)$.

Since

$$| x_n(t) - x_m(t) | \leq C \| x_n - x_m \| = Cd(f_{x_n}, f_{x_m}),$$

$x_n(t)$ is Cauchy, so it converges.

If $\{\tilde{x}_n\}$ is another sequence such that $f_{\tilde{x}_n}$ converges to f, the argument above shows that $x_n(t) - \tilde{x}_n(t) \to 0$; that is, $f_{x_n}(t) - f_{\tilde{x}_n}(t) \to 0$, $\forall t \in T$. Therefore the assignment is well defined.

But does is characterize f completely? That is, if $f(t) \equiv 0$ $\forall t$, does this mean that $f \equiv 0$?

For this, we need an extra condition.

$$\text{If } \{x_n\} \text{ is Cauchy in } X \text{ and } \lim_{n\to\infty} x_n(t) = 0\ \forall t \in T, \text{ then } \| x \| \to 0. \qquad (*)$$

Condition (*) will give us the property we need.

So $(\overline{X}, \langle\,,\rangle)$ is a Hilbert space, it contains $(X, \langle\,,\rangle)$, $\langle\,,\rangle$ induces the inner product on X, moreover, the closure of X is the full space $\overline{X}$.

If $(H, \langle\langle\,,\rangle\rangle)$ is another Hilbert space enjoining the above mentioned properties of $(\overline{X}, \langle\,,\rangle)$, then the identity operator on $(X, \langle\,,\rangle)$ plainly extends to a 1-1, onto unitary operator from $\overline{X}$ to H.

The unique Hilbert space satisfying the above mentioned properties is called the *completion* of $(X, \langle\,,\rangle)$.

Now, going back to our construction, if $(X, \langle\,,\rangle)$ is a vector space of functions on a set T where point evaluations are continuous, then for f_x, set $f_x(t) := x(t)$ and if $f_{x_n} \to f$ in $(\overline{X}, \langle\,,\rangle)$, then we propose to set $f(t) = \lim_{n\to\infty} x_n(t)$.

Since

$$| x_n(t) - x_m(t) | \leq C \| x_n - x_m \| = Cd(f_{x_n}, f_{x_m}),$$

$x_n(t)$ is Cauchy, so it converges.

If $\{\tilde{x}_n\}$ is another sequence such that $f_{\tilde{x}_n}$ converges to f, the argument above shows that $x_n(t) - \tilde{x}_n(t) \to 0$; that is, $f_{x_n}(t) - f_{\tilde{x}_n}(t) \to 0$, $\forall t \in T$. Therefore the assignment is well defined.

But does is characterize f completely? That is, if $f(t) \equiv 0\ \forall t$, does this mean that $f \equiv 0$?

For this, we need an extra condition.

$$\text{If } \{x_n\} \text{ is Cauchy in } X \text{ and } \lim_{n\to\infty} x_n(t) = 0\ \forall t \in T, \text{ then } \| x \| \to 0. \qquad (*)$$

Condition (*) will give us the property we need.

On the other hand if f is completely determined by T then for any Cauchy sequence $\{x_n\} \in X$, $x_n \to x$ in the closure, $x_n(t) \to 0 \; \forall t \implies x(t) \equiv 0 \implies \| x_n \| \to 0$.

So condition (*) is what we seek.

To Summarize

Given an inner product space $(X_0, \langle\,,\rangle_0)$, there exists a unique Hilbert space $(X, \langle\,,\rangle)$ containing a copy of X_0 in the sense that $\exists \imath : X_0 \hookrightarrow X$ one to one linear map that satisfies

- $\langle \imath(x), \imath(y) \rangle = \langle x, y \rangle_0$, and
- $\overline{\imath(X_0)} = X$.

If $(X_0, \langle\,,\rangle_0)$ is a function space on T with continuous point evaluations, there exists a Hilbert function space $(X_1, \langle\,,\rangle_1)$ on T with continuous point evaluations and satisfying the above conditions if and only if

$$\{x_n\}_n \text{ Cauchy in } X_0,\; x_n(t) \to 0 \;\forall t \in T \implies \| x_n \| = \langle x_n, x_n \rangle_0^{1/2} \to 0 \text{ as } n \to \infty \qquad (*)$$

On the other hand if f is completely determined by T then for any Cauchy sequence $\{x_n\} \in X$, $x_n \to x$ in the closure, $x_n(t) \to 0 \; \forall t \implies x(t) \equiv 0 \implies \| x_n \| \to 0$.

So condition (*) is what we seek.

To Summarize

Given an inner product space $(X_0, \langle\, , \rangle_0)$, there exists a unique Hilbert space $(X, \langle\, , \rangle)$ containing a copy of X_0 in the sense that $\exists \imath : X_0 \hookrightarrow X$ one to one linear map that satisfies

- $\langle \imath(x), \imath(y) \rangle = \langle x, y \rangle_0$, and
- $\overline{\imath(X_0)} = X$.

If $(X_0, \langle\, , \rangle_0)$ is a function space on T with continuous point evaluations, there exists a Hilbert function space $(X_1, \langle\, , \rangle_1)$ on T with continuous point evaluations and satisfying the above conditions if and only if

$$\{x_n\}_n \text{ Cauchy in } X_0,\; x_n(t) \to 0 \; \forall t \in T \implies \| x_n \| = \langle x_n, x_n \rangle_0^{1/2} \to 0 \text{ as } n \to \infty \qquad (*)$$

Note that

Since we were interested in the existence of completion of an inner product X, we did not care about the identification $\overline{X}$ in X^*.

Actually, one can show that $\overline{X}$ is in fact X^* as follows:

Take $\sigma \in X^*$. Then $\overline{\sigma}$ is a continuous linear functional on X, hence can be extended to a continuous linear functional on $\overline{X}$. Using the notation of the proof, Riesz Representation Theorem applied to the Hilbert space $(\overline{X}, \langle \, , \rangle)$ gives an element $\eta \in \overline{X}$ such that:

$$\overline{\sigma}(x) = \langle f_x, \eta \rangle = \overline{\langle \eta, f_x \rangle} = \overline{\eta}(x) \quad \Longrightarrow \quad \sigma \in \overline{X}.$$

Moreover, the proof also shows that the norm $\| \, . \, \|$ on X^* is actually a Hilbertian norm, that is, it comes from an inner product on X^*.

In our previous discussions we have represented, from time to time, a given Hilbert space as a space of functions on a set T with the property that point evaluations are continuous. Namely, we have associated elements of a given Hilbert space $(H, \langle\, , \rangle)$ to functions on the **set** H via the rule $H \ni x \leftrightarrow \hat{x}(h) := \langle h, x \rangle$, i.e., as continuous linear functions from H into $\mathbb{C}$ in view of Riesz Representation Theorem. Transporting the inner product to this space of functions, i.e., setting $\langle \hat{x}, \hat{y} \rangle := \langle x, y \rangle \; \forall x, y \in H$, one can view H as a Hilbert space of functions such that point evaluations are continuous.

The last assertion follows immediately from:

$$| \hat{x}(t) | = | \langle t, x \rangle | \leq \| x \| \| t \|, \quad \forall x, t \in H.$$

However, the above realization is not unique. For example, one can also view $\mathbb{R}^n$, or more generally ℓ_2 as a space of functions on $f : \mathbb{N} \to \mathbb{R}$ with $\sum_{n=1}^{\infty} | f(n) |^2 < \infty$, with the inner product:

$$\langle f, g \rangle = \sum_{n=1}^{\infty} f(n) \overline{g(n)},$$

via $\ell_2 \ni \{a_n\} = x \leftrightarrow f_x : f_x(n) = a_n \; \forall n$.

Clearly;

$$| f(n) | \leq \left(\sum_{k=1}^{\infty} | f(k) |^2 \right)^{1/2},$$

so indeed point evaluations are continuous on this function space.

In our previous discussions we have represented, from time to time, a given Hilbert space as a space of functions on a set T with the property that point evaluations are continuous. Namely, we have associated elements of a given Hilbert space $(H, \langle\, , \rangle)$ to functions on the **set** H via the rule $H \ni x \leftrightarrow \hat{x}(h) := \langle h, x \rangle$, i.e., as continuous linear functions from H into $\mathbb{C}$ in view of Riesz Representation Theorem. Transporting the inner product to this space of functions, i.e., setting $\langle \hat{x}, \hat{y} \rangle := \langle x, y \rangle \; \forall x, y \in H$, one can view H as a Hilbert space of functions such that point evaluations are continuous.

The last assertion follows immediately from:

$$| \hat{x}(t) | = | \langle t, x \rangle | \leq \| x \| \| t \|, \quad \forall x, t \in H.$$

However, the above realization is not unique. For example, one can also view $\mathbb{R}^n$, or more generally ℓ_2 as a space of functions on $f : \mathbb{N} \to \mathbb{R}$ with $\sum_{n=1}^{\infty} | f(n) |^2 < \infty$, with the inner product:

$$\langle f, g \rangle = \sum_{n=1}^{\infty} f(n)\overline{g(n)},$$

via $\ell_2 \ni \{a_n\} = x \leftrightarrow f_x : f_x(n) = a_n \; \forall n$.

Clearly;

$$| f(n) | \leq \left(\sum_{k=1}^{\infty} | f(k) |^2 \right)^{1/2},$$

so indeed point evaluations are continuous on this function space.

Naturally, it is desirable to represent a given Hilbert space as a function space on a small set.

On the other hand, some important Hilbert spaces that occur in nature are given as function spaces with continuous point evaluations.

Example 1

Let X be the vector space of all infinitely differentiable real valued functions which vanish outside of a finite interval and let

$$\langle f, g\rangle := \int_{-\infty}^{\infty}(fg)(t)dt + \int_{-\infty}^{\infty}(f'g')(t)dt.$$

Then $(X, \langle\,,\rangle)$ becomes an inner product space.

For a point $t_0 \in \mathbb{R}$ and $f \in X$, by the Fundamental Theorem of Calculus,

$$\begin{aligned} |\, f(t_0)\,|^2 &= \left|\int_{-\infty}^{t_0}(f^2)'dt\right| = \left|2\int_{-\infty}^{t_0} f'fdt\right| \\ &\leq 2\Big(\int_{-\infty}^{\infty} |\, f'\,|^2\, dt\Big)^{1/2}\Big(\int_{-\infty}^{\infty} |\, f\,|^2\, dt\Big)^{1/2} \\ &\leq \int_{-\infty}^{\infty} |\, f'\,|^2\, dt + \int_{-\infty}^{\infty} |\, f\,|^2\, dt. \end{aligned}$$

Note that, we have used the inequality $2AB \leq A^2 + B^2$ for positive real numbers A and B.

So, $|\, f(t_0)\,| \leq \langle f, f\rangle^{1/2}$, hence point evaluations are continuous on X.

Naturally, it is desirable to represent a given Hilbert space as a function space on a small set.

On the other hand, some important Hilbert spaces that occur in nature are given as function spaces with continuous point evaluations.

Example 1

Let X be the vector space of all infinitely differentiable real valued functions which vanish outside of a finite interval and let

$$\langle f, g\rangle := \int_{-\infty}^{\infty} (fg)(t)dt + \int_{-\infty}^{\infty} (f'g')(t)dt.$$

Then $(X, \langle\, ,\rangle)$ becomes an inner product space.

For a point $t_0 \in \mathbb{R}$ and $f \in X$, by the Fundamental Theorem of Calculus,

$$\begin{aligned} \mid f(t_0) \mid^2 &= \left|\int_{-\infty}^{t_0} (f^2)' dt\right| = \left|2\int_{-\infty}^{t_0} f' f dt\right| \\ &\leq 2\Big(\int_{-\infty}^{\infty} \mid f' \mid^2 dt\Big)^{1/2}\Big(\int_{-\infty}^{\infty} \mid f \mid^2 dt\Big)^{1/2} \\ &\leq \int_{-\infty}^{\infty} \mid f' \mid^2 dt + \int_{-\infty}^{\infty} \mid f \mid^2 dt. \end{aligned}$$

Note that, we have used the inequality $2AB \leq A^2 + B^2$ for positive real numbers A and B.

So, $\mid f(t_0) \mid \leq \langle f, f\rangle^{1/2}$, hence point evaluations are continuous on X.

Example 1 (Cont.)

We would like to draw attention to two points in this context:

1) The first term in the above inner product is itself an inner product on X, but the point evaluations are not necessarily continuous on this inner product space, as we have observed in a previous example.

2) $(X, \langle\, , \rangle)$ is not complete. However, it can be shown that it satisfies the condition (*) for having a completion consisting of certain continuous functions on $\mathbb{R}$ with continuous point evaluations.

Hence, the completion of $C_c^\infty(\mathbb{R})$, $W(\mathbb{R})$, is a Hilbert space of functions on $\mathbb{R}$ with continuous point evaluations.

As a matter of fact, $W(\mathbb{R})$ consists of continuous functions f on $\mathbb{R}$, differentiable at "most" of the points of $\mathbb{R}$ and $\int_{-\infty}^{\infty} \mid f(t) \mid^2 dt < \infty$, $\int_{-\infty}^{\infty} \mid f'(t) \mid^2 dt < \infty$ with a "reasonable" interpretation of the second integral.

Example 2

Let $H^2(\mathbb{D})$ denote the vector space of all analytic functions on the unit disc $\mathbb{D} \subseteq \mathbb{C}$ whose Taylor coefficients are in ℓ_2. In other words,

$$H^2(\mathbb{D}) := \{f(z) = \sum_{n=1}^{\infty} c_n z^n \text{ on } \mathbb{D} \text{ with } \sum_{n=1}^{\infty} \mid c_n \mid^2 < \infty\}.$$

On $H^2(\mathbb{D})$ we put the inner product:

$$\langle f, g \rangle := \sum_{n=1}^{\infty} c_n \overline{d_n}, \quad f(z) = \sum_{n=1}^{\infty} c_n z^n, \quad g(z) = \sum_{n=1}^{\infty} d_n z^n \in H^2(\mathbb{D}).$$

Example 1 (Cont.)

We would like to draw attention to two points in this context:

1) The first term in the above inner product is itself an inner product on X, but the point evaluations are not necessarily continuous on this inner product space, as we have observed in a previous example.
2) $(X, \langle \, , \rangle)$ is not complete. However, it can be shown that it satisfies the condition (*) for having a completion consisting of certain continuous functions on $\mathbb{R}$ with continuous point evaluations.

Hence, the completion of $C_c^\infty(\mathbb{R})$, $W(\mathbb{R})$, is a Hilbert space of functions on $\mathbb{R}$ with continuous point evaluations.
As a matter of fact, $W(\mathbb{R})$ consists of continuous functions f on $\mathbb{R}$, differentiable at "most" of the points of $\mathbb{R}$ and $\int_{-\infty}^{\infty} \mid f(t) \mid^2 dt < \infty$, $\int_{-\infty}^{\infty} \mid f'(t) \mid^2 dt < \infty$ with a "reasonable" interpretation of the second integral.

Example 2

Let $H^2(\mathbb{D})$ denote the vector space of all analytic functions on the unit disc $\mathbb{D} \subseteq \mathbb{C}$ whose Taylor coefficients are in ℓ_2. In other words,

$$H^2(\mathbb{D}) := \{f(z) = \sum_{n=1}^{\infty} c_n z^n \text{ on } \mathbb{D} \text{ with } \sum_{n=1}^{\infty} \mid c_n \mid^2 < \infty\}.$$

On $H^2(\mathbb{D})$ we put the inner product:

$$\langle f, g \rangle := \sum_{n=1}^{\infty} c_n \overline{d_n}, \quad f(z) = \sum_{n=1}^{\infty} c_n z^n, \quad g(z) = \sum_{n=1}^{\infty} d_n z^n \in H^2(\mathbb{D}).$$

Example 2 (Cont.)

Clearly $\langle\,,\rangle$ defines an inner product on $H^2(\mathbb{D})$ and it makes it a Hilbert space, basically because ℓ_2 is complete.

Note that, if a sequence $\{f_n\}_n$, $f_n = \sum_{k=1}^{\infty} a_k^n z^k$ is Cauchy, then $x_n := \{a_k^n\}_{k=1}^{\infty}$, $n = 1, 2, \cdots$ is a Cauchy sequence in ℓ_2, so converges to some $x = \{a_k\}_{k=1}^{\infty} \in \ell_2$.

Now, $f(z) = \sum_{k=1}^{\infty} a_k z^n$ defines a function on $\mathbb{D}$ since on each subdisc Δ_r, where $\Delta_r := \{z :| z |< r\}$, $r < 1$:

$$\sum_{k=1}^{\infty} | a_k || z^n |\leq \sum_{k=1}^{n} | a_k | r^n \leq \Big(\sum_{k=1}^{\infty} | a_k |^2 \Big)^{1/2} \Big(\frac{1}{1 - r^2} \Big)^{1/2},$$

and this function is analytic since it is the uniform limit of analytic polynomials $\sum_{k=1}^{N} a_k z^n$ on each subdisc $\Delta_r, r < 1$.

For a $w \in \mathbb{D}$ and $f \in H^2(\mathbb{D})$,

$$| f(w) |\leq \sum_{n=1}^{\infty} | a_n || w |^n \leq \Big(\frac{1}{1-| w |^2} \Big)^{1/2} \| f \| .$$

So $H^2(\mathbb{D})$ is a Hilbert space of functions on the unit disc $\mathbb{D}$ with continuous point evaluations.

Example 2 (Cont.)

Clearly $\langle\, , \rangle$ defines an inner product on $H^2(\mathbb{D})$ and it makes it a Hilbert space, basically because ℓ_2 is complete.

Note that, if a sequence $\{f_n\}_n$, $f_n = \sum_{k=1}^{\infty} a_k^n z^k$ is Cauchy, then $x_n := \{a_k^n\}_{k=1}^{\infty}$, $n = 1, 2, \cdots$ is a Cauchy sequence in ℓ_2, so converges to some $x = \{a_k\}_{k=1}^{\infty} \in \ell_2$.

Now, $f(z) = \sum_{k=1}^{\infty} a_k z^n$ defines a function on $\mathbb{D}$ since on each subdisc Δ_r, where $\Delta_r := \{z :| z |< r\}$, $r < 1$:

$$\sum_{k=1}^{\infty} | a_k || z^n | \leq \sum_{k=1}^{n} | a_k | r^n \leq \left(\sum_{k=1}^{\infty} | a_k |^2 \right)^{1/2} \left(\frac{1}{1 - r^2} \right)^{1/2},$$

and this function is analytic since it is the uniform limit of analytic polynomials $\sum_{k=1}^{N} a_k z^n$ on each subdisc $\Delta_r, r < 1$.

For a $w \in \mathbb{D}$ and $f \in H^2(\mathbb{D})$,

$$| f(w) | \leq \sum_{n=1}^{\infty} | a_n || w |^n \leq \left(\frac{1}{1 - | w |^2} \right)^{1/2} \| f \| .$$

So $H^2(\mathbb{D})$ is a Hilbert space of functions on the unit disc $\mathbb{D}$ with continuous point evaluations.

Example 3

Suppose T is any set (could be a set of humans for example), and suppose we somehow fabricate a map ϕ from T to a Hilbert space $(H, \langle\, , \rangle)$ not necessarily a RKHS.

Elaborating on the comment above, as an example, one can use the assignment of their weight/height/birth year to a human in the set T, so ϕ from T to $\mathbb{R}^3$ with the usual inner product, becomes a function.

This map induces a RKHS of functions on T by first considering:

$$\mathcal{H} := \overline{span_{t \in T} \phi(t)} \subset H,$$

and forming:

$$\tilde{h}(t) := \langle h, \phi(t) \rangle, \text{ for } h \in \mathcal{H},$$

with the inner product:

$$\langle \tilde{h}_1, \tilde{h}_2 \rangle := \langle h_1, h_2 \rangle.$$

That is, we think of elements of $\mathcal{H}$ as functions on H and restrict them to the image of ϕ. Note that $\tilde{h}(t) \equiv 0$ implies $h \equiv 0$ since $h \in \mathcal{H}$.

Example 3

Suppose T is any set (could be a set of humans for example), and suppose we somehow fabricate a map ϕ from T to a Hilbert space $(H, \langle\, , \rangle)$ not necessarily a RKHS.

Elaborating on the comment above, as an example, one can use the assignment of their weight/height/birth year to a human in the set T, so ϕ from T to $\mathbb{R}^3$ with the usual inner product, becomes a function.

This map induces a RKHS of functions on T by first considering:

$$\mathcal{H} := \overline{span_{t \in T} \phi(t)} \subset H,$$

and forming:

$$\tilde{h}(t) := \langle h, \phi(t) \rangle, \text{ for } h \in \mathcal{H},$$

with the inner product:

$$\langle \tilde{h}_1, \tilde{h}_2 \rangle := \langle h_1, h_2 \rangle.$$

That is, we think of elements of $\mathcal{H}$ as functions on H and restrict them to the image of ϕ. Note that $\tilde{h}(t) \equiv 0$ implies $h \equiv 0$ since $h \in \mathcal{H}$.

Example 3 (Cont.)

Another way of visualizing this example is by forming the function space

$$\mathcal{F} = \{f : T \to \mathbb{C} \mid \exists h \in H : f(t) = \langle h, \phi(t) \rangle \; \forall t \in T\}$$

and putting on $\mathcal{F}$ the norm

$$\| f \| := \inf_{h \in H, f(t) = \langle h, \phi(t) \rangle} \| h \|,$$

where the last norm is the norm in the Hilbert space H.

Note that with the notation above,

$$h \in \mathcal{H}^{\perp} \iff \langle h, \phi(t) \rangle = 0 \; \forall t \in T.$$

It follows that for every $f \in \mathcal{F}$ there exits a unique $h_f \in \mathcal{H}$ such that $f(t) = \langle h_f, \phi(t) \rangle \; \forall t \in T$, and if $f(t) = \langle h, \phi(t) \rangle \; \forall t \in T$, then $h = h_1 + h_f$ with $h_1 \in \mathcal{H}^{\perp}$ so $\| h \| \geq \| h_f \|$.

It follows that $\| f \|^2 = \langle h_f, h_f \rangle$, hence the norm on $\mathcal{F}$ is coming from an inner product in view of the polarization identity and the assignment $f \longleftrightarrow h_f$ is a unitary isomorphism between the Hilbert spaces $\mathcal{H}$ and $\mathcal{F}$.

Example 3 (Cont.)

Another way of visualizing this example is by forming the function space

$$\mathcal{F} = \{f : T \to \mathbb{C} \,|\, \exists h \in H : f(t) = \langle h, \phi(t) \rangle \; \forall t \in T\}$$

and putting on $\mathcal{F}$ the norm

$$\| f \| := \inf_{h \in H, f(t) = \langle h, \phi(t) \rangle} \| h \|,$$

where the last norm is the norm in the Hilbert space H.

Note that with the notation above,

$$h \in \mathcal{H}^{\perp} \iff \langle h, \phi(t) \rangle = 0 \; \forall t \in T.$$

It follows that for every $f \in \mathcal{F}$ there exits a unique $h_f \in \mathcal{H}$ such that $f(t) = \langle h_f, \phi(t) \rangle \; \forall t \in T$, and if $f(t) = \langle h, \phi(t) \rangle \; \forall t \in T$, then $h = h_1 + h_f$ with $h_1 \in \mathcal{H}^{\perp}$ so $\| h \| \geq \| h_f \|$.

It follows that $\| f \|^2 = \langle h_f, h_f \rangle$, hence the norm on $\mathcal{F}$ is coming from an inner product in view of the polarization identity and the assignment $f \longleftrightarrow h_f$ is a unitary isomorphism between the Hilbert spaces $\mathcal{H}$ and $\mathcal{F}$.

In the Hilbert spaces $(H, \langle\,,\rangle)$ of functions, on a set T, for which point evaluations are continuous, like the examples given above, H possesses a collection of distinguished elements $\{\mathbb{k}_t\}_{t\in T}$ defined by:

$$\boxed{\langle x, \mathbb{k}_t\rangle := x(t), \quad \forall x \in H.} \tag{4}$$

Note that such vectors exist in view of Riesz Representation Theorem since point evaluations are continuous and are uniquely determined by the given point of T.
One can think of points $h \in \mathcal{H}$ as indexed by elements of T as $\{x(t)\}_{t\in T}$. The importance of the elements $\mathbb{k}_t, t \in T$ is that they give "coordinates" $\{x(t)\}_{t\in T}$ of an $x \in H$ by using the inner product on H via equality (4).
Hence, for example, if one theoretically knows that a sequence $\{x_n\}$ converges, the knowledge of these distinguished vectors will allow us to compute the "coordinates" of the limit vector x via
$x(t) = \lim\langle x, \mathbb{k}_t\rangle, \quad \forall t \in T.$

Definition

Let $(H, \langle\,,\rangle)$ be a Hilbert space of functions on T such that the point evaluations are continuous. Let $\{\mathbb{k}_t\}_{t\in T}$ be the vectors of H defined as above. One calls the scalar valued function defined on $T \times T$ via:

$$K(t,s) := \langle \mathbb{k}_s, \mathbb{k}_t\rangle = \mathbb{k}_s(t) = \overline{\mathbb{k}_t(s)} = \overline{\langle \mathbb{k}_t, \mathbb{k}_s\rangle},$$

the **kernel** of $(H, \langle\,,\rangle)$.

This kernel is reproducing in the sense that it captures the "coordinates" of $x \in H$ via:

$$x(s) = \langle x, \mathbb{k}_s\rangle = \langle x, K(s,.)\rangle, \quad \forall s,$$

since the function $t \mapsto K(s,t)$ is just an element $\mathbb{k}_s$ in H.

In the Hilbert spaces $(H, \langle\,,\rangle)$ of functions, on a set T, for which point evaluations are continuous, like the examples given above, H possesses a collection of distinguished elements $\{\Bbbk_t\}_{t\in T}$ defined by:

$$\langle x, \Bbbk_t \rangle := x(t), \quad \forall x \in H. \tag{4}$$

Note that such vectors exist in view of Riesz Representation Theorem since point evaluations are continuous and are uniquely determined by the given point of T.
One can think of points $h \in \mathcal{H}$ as indexed by elements of T as $\{x(t)\}_{t\in T}$. The importance of the elements $\Bbbk_t, t \in T$ is that they give "coordinates" $\{x(t)\}_{t\in T}$ of an $x \in H$ by using the inner product on H via equality (4).
Hence, for example, if one theoretically knows that a sequence $\{x_n\}$ converges, the knowledge of these distinguished vectors will allow us to compute the "coordinates" of the limit vector x via
$x(t) = \lim\langle x, \Bbbk_t\rangle, \quad \forall t \in T.$

Definition

Let $(H, \langle\,,\rangle)$ be a Hilbert space of functions on T such that the point evaluations are continuous. Let $\{\Bbbk_t\}_{t\in T}$ be the vectors of H defined as above. One calls the scalar valued function defined on $T \times T$ via:

$$K(t,s) := \langle \Bbbk_s, \Bbbk_t\rangle = \Bbbk_s(t) = \overline{\Bbbk_t(s)} = \overline{\langle \Bbbk_t, \Bbbk_s\rangle},$$

the **kernel** of $(H, \langle\,,\rangle)$.

This kernel is reproducing in the sense that it captures the "coordinates" of $x \in H$ via:

$$x(s) = \langle x, \Bbbk_s\rangle = \langle x, K(s,.)\rangle, \quad \forall s,$$

since the function $t \mapsto K(s,t)$ is just an element $\Bbbk_s$ in H.

Theorem

Let K, defined on $T \times T$ as above, be the kernel function of a Hilbert space $(H, \langle\,,\rangle)$. We have:

1. $K(t,s) = \overline{K(s,t)}, \quad \forall s,t \in T,$
2. *For* $(\lambda_1, \cdots, \lambda_N) \in \mathbb{C}^N$; $\sum_{i,j} \lambda_i \overline{\lambda_j} K(t_i, t_j) \geq 0, \quad \forall N \in \mathbb{N}, (t_1, \cdots, t_N) \in T^N$.

Proof.

The first property is clear.

To see the second, choose $(\lambda_1, \cdots, \lambda_N) \in \mathbb{C}^N$ and $(t_1, \cdots, t_N) \in T^N$;

$$0 \leq \left\langle \sum_{i=1}^{N} \lambda_i \mathbb{k}_i, \sum_{i=1}^{N} \lambda_i \mathbb{k}_i \right\rangle = \sum_{i,j=1}^{N} \lambda_i \overline{\lambda_j} \langle \mathbb{k}_i, \mathbb{k}_j \rangle = \sum_{i,j=1}^{N} \lambda_i \overline{\lambda_j} K(t_i, t_j).$$

□

The second condition is usually referred to as **positiveness** of K since it is just the condition that the matrix $\left(K(t_i, t_j)\right)_{i,j=1}^{N}$, $N \in \mathbb{N}$ is a positive matrix.

Theorem

Let K, defined on $T \times T$ as above, be the kernel function of a Hilbert space $(H, \langle\, , \rangle)$. We have:

1. $K(t,s) = \overline{K(s,t)}, \quad \forall s, t \in T,$
2. *For* $(\lambda_1, \cdots, \lambda_N) \in \mathbb{C}^N$; $\sum_{i,j} \lambda_i \overline{\lambda_j} K(t_i, t_j) \geq 0, \quad \forall N \in \mathbb{N}, (t_1, \cdots, t_N) \in T^N$.

Proof.

The first property is clear.

To see the second, choose $(\lambda_1, \cdots, \lambda_N) \in \mathbb{C}^N$ and $(t_1, \cdots, t_N) \in T^N$;

$$0 \leq \left\langle \sum_{i=1}^{N} \lambda_i \mathbb{k}_i, \sum_{i=1}^{N} \lambda_i \mathbb{k}_i \right\rangle = \sum_{i,j=1}^{N} \lambda_i \overline{\lambda_j} \langle \mathbb{k}_i, \mathbb{k}_j \rangle = \sum_{i,j=1}^{N} \lambda_i \overline{\lambda_j} K(t_i, t_j).$$

□

The second condition is usually referred to as **positiveness** of K since it is just the condition that the matrix $\left(K(t_i, t_j)\right)_{i,j=1}^{N}$, $N \in \mathbb{N}$ is a positive matrix.

It is time to give such Hilbert function spaces a name:

Definition

A **Reproducing Kernel Hilbert Space (RKHS)** is a Hilbert space $(H, \langle\, , \rangle)$ of functions on a set T such that all the point evaluations are continuous.

In the definition we have suppressed the kernel function, however we will see later that this scalar valued function completely determines the Hilbert space. We revisit the examples given above.

Example 1*

Let

$$\mathcal{X}(s,t) = e^{-|t-s|} = \begin{cases} e^{-t+s} & \text{if } t > s, \\ e^{t-s} & \text{if } t \leq s \end{cases} \quad s, t \in \mathbb{R}.$$

Observe that

$$\frac{\partial \mathcal{X}(s,t)}{\partial t} = \begin{cases} -e^{-t+s} & \text{if } t > s, \\ e^{t-s} & \text{if } t < s \end{cases} \quad s, t \in \mathbb{R}.$$

So $\mathcal{X}_s(t) := \mathcal{X}(s,t)$ is differentiable except at the point s and $\int_{-\infty}^{\infty} \left| \frac{\partial \mathcal{X}}{\partial t}(t) \right|^2 dt$ is finite if we interpret the integral as:

$$\int_{-\infty}^{\infty} \left| \frac{\partial \mathcal{X}_s}{\partial t}(t) \right|^2 dt = \int_{-\infty}^{s} \left| \frac{\partial \mathcal{X}_s}{\partial t}(t) \right|^2 dt + \int_{s}^{\infty} \left| \frac{\partial \mathcal{X}_s}{\partial t}(t) \right|^2 dt.$$

It follows that $\mathcal{X}_s \in W(\mathbb{R})$ in view of our previous discussion.

It is time to give such Hilbert function spaces a name:

Definition

A **Reproducing Kernel Hilbert Space (RKHS)** is a Hilbert space $(H, \langle\,,\rangle)$ of functions on a set T such that all the point evaluations are continuous.

In the definition we have suppressed the kernel function, however we will see later that this scalar valued function completely determines the Hilbert space. We revisit the examples given above.

Example 1*

Let

$$\mathcal{X}(s,t) = e^{-|t-s|} = \begin{cases} e^{-t+s} & \text{if } t > s, \\ e^{t-s} & \text{if } t \leq s \end{cases} \quad s,t \in \mathbb{R}.$$

Observe that

$$\frac{\partial \mathcal{X}(s,t)}{\partial t} = \begin{cases} -e^{-t+s} & \text{if } t > s, \\ e^{t-s} & \text{if } t < s \end{cases} \quad s,t \in \mathbb{R}.$$

So $\mathcal{X}_s(t) := \mathcal{X}(s,t)$ is differentiable except at the point s and $\int_{-\infty}^{\infty} \left|\frac{\partial \mathcal{X}}{\partial t}(t)\right|^2 dt$ is finite if we interpret the integral as:

$$\int_{-\infty}^{\infty} \left|\frac{\partial \mathcal{X}_s}{\partial t}(t)\right|^2 dt = \int_{-\infty}^{s} \left|\frac{\partial \mathcal{X}_s}{\partial t}(t)\right|^2 dt + \int_{s}^{\infty} \left|\frac{\partial \mathcal{X}_s}{\partial t}(t)\right|^2 dt.$$

It follows that $\mathcal{X}_s \in W(\mathbb{R})$ in view of our previous discussion.

Example 1* (Cont.)

For $f \in C_c^\infty(\mathbb{R})$, we compute:

$$\int_{-\infty}^{\infty} f(t)\mathcal{X}_s(t)dt = \int_{-\infty}^{s} f(t)e^{t-s}dt + \int_{s}^{\infty} f(t)e^{-t+s}dt$$

$$\begin{aligned}\int_{-\infty}^{\infty} f'(t)\mathcal{X}_s'(t)dt &= \int_{-\infty}^{s} f'(t)e^{t-s}dt - \int_{s}^{\infty} f'(t)e^{-t+s}dt \\ &= -\int_{-\infty}^{s} f(t)e^{t-s}dt + f(s) - \left(\int_{s}^{\infty} f(t)e^{-t+s}dt - f(s)\right) \\ &= -\int_{-\infty}^{s} f(t)e^{t-s}dt - \int_{s}^{\infty} f(t)e^{-t+s}dt + 2f(s),\end{aligned}$$

So

$$\int_{-\infty}^{\infty} f(t)\mathcal{X}_s(t)dt + \int_{-\infty}^{\infty} f'(t)\mathcal{X}_s'(t)dt = 2f(s).$$

So $\langle f, \frac{1}{2}\mathcal{X}_s\rangle = f(s)$ for $f \in C_c^\infty(\mathbb{R})$, hence for $f \in W$; since $\overline{C_c^\infty(\mathbb{R})} = W$ and point evaluations are continuous on W.

Thus the kernel on W is the function:

$$K : \mathbb{R}^2 \to \mathbb{R} \quad K(s,t) = \langle \frac{1}{2}\mathcal{X}_t, \frac{1}{2}\mathcal{X}_s\rangle = \frac{1}{4}e^{-|t-s|}.$$

Example 1* (Cont.)

For $f \in C_c^\infty(\mathbb{R})$, we compute:

$$\int_{-\infty}^{\infty} f(t)\mathcal{X}_s(t)dt = \int_{-\infty}^{s} f(t)e^{t-s}dt + \int_{s}^{\infty} f(t)e^{-t+s}dt$$

$$\int_{-\infty}^{\infty} f'(t)\mathcal{X}_s'(t)dt = \int_{-\infty}^{s} f'(t)e^{t-s}dt - \int_{s}^{\infty} f'(t)e^{-t+s}dt$$

$$= -\int_{-\infty}^{s} f(t)e^{t-s}dt + f(s) - \left(\int_{s}^{\infty} f(t)e^{-t+s}dt - f(s)\right)$$

$$= -\int_{-\infty}^{s} f(t)e^{t-s}dt - \int_{s}^{\infty} f(t)e^{-t+s}dt + 2f(s),$$

So

$$\int_{-\infty}^{\infty} f(t)\mathcal{X}_s(t)dt + \int_{-\infty}^{\infty} f'(t)\mathcal{X}_s'(t)dt = 2f(s).$$

So $\langle f, \frac{1}{2}\mathcal{X}_s\rangle = f(s)$ for $f \in C_c^\infty(\mathbb{R})$, hence for $f \in W$; since $\overline{C_c^\infty(\mathbb{R})} = W$ and point evaluations are continuous on W.

Thus the kernel on W is the function:

$$K : \mathbb{R}^2 \to \mathbb{R} \quad K(s,t) = \langle \frac{1}{2}\mathcal{X}_t, \frac{1}{2}\mathcal{X}_s\rangle = \frac{1}{4}e^{-|t-s|}.$$

Before we proceed further, a simple observation is in order.

Suppose $H \subseteq \mathbb{C}^T$ is a RKHS with kernel K that is separable, i.e., it contains countable elements $f_n, n = 1, 2, \cdots$ such that $\overline{span\{f_n\}} = H$. Then any orthonormal basis of H is countable.
Suppose $\{e_n\}$ is any such orthonormal basis for H, then $\forall t \in T$ consider the expansion in H,

$$\mathbb{k}_t(s) = \sum_{n=1}^{\infty} \langle \mathbb{k}_t, e_n \rangle e_n(s) = \sum_{n=1}^{\infty} \overline{e_n(t)} e_n = \lim_{N\to\infty} \sum_{n=1}^{N} \overline{e_n(t)} e_n \text{ in } H.$$

Since point evaluations are continuous on H,

$$\mathbb{k}_t(s) = \lim_{N\to\infty} \mathbb{k}_t\Big(\sum_{n=1}^{N} \overline{e_n(t)} e_n(s) \Big) = \lim_{N\to\infty} \sum_{n=1}^{N} e_n(s) \overline{e_n(t)}$$
$$= \sum_{n=1}^{\infty} e_n(s) \overline{e_n(t)},$$

where the last convergence is in $\mathbb{C}$.

Hence the kernel of H can be computed as

$$K(t, s) = \mathbb{k}_t(s) = \sum_{n=1}^{\infty} e_n(s) \overline{e_n(t)}.$$

Since the right hand side of the equation depends only on H, any choice of orthonormal basis can be used to compute the kernel of H.

Before we proceed further, a simple observation is in order.

Suppose $H \subseteq \mathbb{C}^T$ is a RKHS with kernel K that is separable, i.e., it contains countable elements $f_n, n = 1, 2, \cdots$ such that $\overline{span\{f_n\}} = H$. Then any orthonormal basis of H is countable. Suppose $\{e_n\}$ is any such orthonormal basis for H, then $\forall t \in T$ consider the expansion in H,

$$\mathbb{k}_t(s) = \sum_{n=1}^{\infty} \langle \mathbb{k}_t, e_n \rangle e_n(s) = \sum_{n=1}^{\infty} \overline{e_n(t)} e_n = \lim_{N\to\infty} \sum_{n=1}^{N} \overline{e_n(t)} e_n \text{ in } H.$$

Since point evaluations are continuous on H,

$$\begin{aligned}\mathbb{k}_t(s) &= \lim_{N\to\infty} \mathbb{k}_t\Big(\sum_{n=1}^{N} \overline{e_n(t)} e_n(s) \Big) = \lim_{N\to\infty} \sum_{n=1}^{N} e_n(s)\overline{e_n(t)} \\ &= \sum_{n=1}^{\infty} e_n(s)\overline{e_n(t)},\end{aligned}$$

where the last convergence is in $\mathbb{C}$.

Hence the kernel of H can be computed as

$$K(t,s) = \mathbb{k}_t(s) = \sum_{n=1}^{\infty} e_n(s)\overline{e_n(t)}.$$

Since the right hand side of the equation depends only on H, any choice of orthonormal basis can be used to compute the kernel of H.

Example 2*

Let f be an analytic function on the unit disc and consider its Taylor series expansion:

$$f = \sum_{n=0}^{\infty} c_n z^n = \sum_{n=0}^{\infty} \langle f, z^n \rangle z^n.$$

The last expression in the right hand side comes directly from the definition of the inner product on $H^2(\mathbb{D})$. Moreover,

$$\left\| \sum_{n=1}^{N} \langle f, z^n \rangle z^n - f \right\|^2 = \sum_{n>N} \mid c_n \mid^2 \to 0 \text{ as } N \to \infty.$$

So the series $\sum_{n=1}^{\infty} \langle f, z^n \rangle z^n$ not only converges uniformly on each disc $\Delta_r = \{z :\mid z \mid < r\}$, but also converges to f in $H^2(\mathbb{D})$, and $\langle z^n, z^m \rangle = \delta_{n,m}$, $\forall n, m \in \mathbb{N}$.

It follows that $\{z^n\}_{n=0}^{\infty}$ is an orthonormal basis in $H^2(\mathbb{D})$.

So the kernel of $H^2(\mathbb{D})$ is:

$$K(\zeta, \eta) = \sum_{n=0}^{\infty} z^n(\eta)\overline{z^n(\zeta)} = \sum_{n=0}^{\infty} (\eta\bar{\zeta})^n = \frac{1}{1 - \eta\bar{\zeta}}.$$

Now we wish to relate this kernel function, obtained by functional analytic considerations to a well known formula of Complex Analysis.

Example 2*

Let f be an analytic function on the unit disc and consider its Taylor series expansion:

$$f = \sum_{n=0}^{\infty} c_n z^n = \sum_{n=0}^{\infty} \langle f, z^n \rangle z^n.$$

The last expression in the right hand side comes directly from the definition of the inner product on $H^2(\mathbb{D})$. Moreover,

$$\left\| \sum_{n=1}^{N} \langle f, z^n \rangle z^n - f \right\|^2 = \sum_{n>N} | c_n |^2 \to 0 \text{ as } N \to \infty.$$

So the series $\sum_{n=1}^{\infty} \langle f, z^n \rangle z^n$ not only converges uniformly on each disc $\Delta_r = \{z :| z |< r\}$, but also converges to f in $H^2(\mathbb{D})$, and $\langle z^n, z^m \rangle = \delta_{n,m}$, $\forall n, m \in \mathbb{N}$.

It follows that $\{z^n\}_{n=0}^{\infty}$ is an orthonormal basis in $H^2(\mathbb{D})$.

So the kernel of $H^2(\mathbb{D})$ is:

$$K(\zeta, \eta) = \sum_{n=0}^{\infty} z^n(\eta)\overline{z^n(\zeta)} = \sum_{n=0}^{\infty} (\eta\bar{\zeta})^n = \frac{1}{1 - \eta\bar{\zeta}}.$$

Now we wish to relate this kernel function, obtained by functional analytic considerations to a well known formula of Complex Analysis.

Example 2* (Cont.)

Let f, g be analytic functions on an open disc containing $\overline{\mathbb{D}} := \{z :| z | \leq 1\}$. Say $f(z) = \sum_{n=0}^{\infty} c_n z^n$ and $g(z) = \sum_{n=0}^{\infty} d_n z^n$. Since:

$$f\overline{g}(e^{i\theta}) = \sum_{n,m=0}^{\infty} c_n \overline{d_m} e^{i(n-m)\theta}$$

on the unit circle and since the series converge uniformly on the unit circle, we have

$$\int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})} d\theta = 2\pi \sum_{n=0}^{\infty} c_n \overline{d_n}.$$

In particular, such functions are in $H^2(\mathbb{D})$ and the above expression represents the inner product of two such functions as an integral.

Note that for any $z \in \mathbb{D}$,

$$\mathbb{k}_z(w) = \frac{1}{1 - w\overline{z}}$$

is an analytic function near the closed unit disc. Therefore for an f that is analytic near the closed unit disc and a point $z_0 = re^{i\psi}$ in the unit disc,

$$f(z_0) = \langle f, \mathbb{k}_{z_0} \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta})\overline{\mathbb{k}_{z_0}(e^{i\theta})} d\theta = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) \frac{d\theta}{1 - z_0 e^{-i\theta}}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{f(e^{i\theta})}{1 - z_0 e^{-i\theta}} \cdot \frac{ie^{i\theta} d\theta}{ie^{i\theta}} = \frac{1}{2\pi i} \int_\Gamma \frac{f(w)}{w - z_0} dw \quad (\text{take } w = e^{i\theta},\ dw = ie^{i\theta} d\theta)$$

where $\Gamma = \partial\mathbb{D}$.

This formula is the classical Cauchy Integral Formula of Complex Analysis.

Example 2* (Cont.)

Let f, g be analytic functions on an open disc containing $\overline{\mathbb{D}} := \{z :| z |\leq 1\}$. Say $f(z) = \sum_{n=0}^{\infty} c_n z^n$ and $g(z) = \sum_{n=0}^{\infty} d_n z^n$. Since:

$$f\overline{g}(e^{i\theta}) = \sum_{n,m=0}^{\infty} c_n \overline{d_m} e^{i(n-m)\theta}$$

on the unit circle and since the series converge uniformly on the unit circle, we have

$$\int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}d\theta = 2\pi \sum_{n=0}^{\infty} c_n \overline{d_n}.$$

In particular, such functions are in $H^2(\mathbb{D})$ and the above expression represents the inner product of two such functions as an integral.

Note that for any $z \in \mathbb{D}$,

$$\mathbb{k}_z(w) = \frac{1}{1 - w\overline{z}}$$

is an analytic function near the closed unit disc. Therefore for an f that is analytic near the closed unit disc and a point $z_0 = re^{i\psi}$ in the unit disc,

$$f(z_0) = \langle f, \mathbb{k}_{z_0} \rangle = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\overline{\mathbb{k}_{z_0}(e^{i\theta})}d\theta = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\frac{d\theta}{1 - z_0 e^{-i\theta}}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{i\theta})}{1 - z_0 e^{-i\theta}} \cdot \frac{ie^{i\theta}d\theta}{ie^{i\theta}} = \frac{1}{2\pi i}\int_\Gamma \frac{f(w)}{w - z_0}dw \quad (\text{take } w = e^{i\theta},\ dw = ie^{i\theta}d\theta)$$

where $\Gamma = \partial\mathbb{D}$.
This formula is the classical Cauchy Integral Formula of Complex Analysis.

Example 3*

In this example, since for $h \in H$, $\tilde{h}(t) = \langle h, \phi(t) \rangle = \langle \tilde{h}, \tilde{\phi}(t) \rangle$, the kernel function is plainly

$$K(s,t) = \mathbb{k}_t(s) = \langle \tilde{\phi}(t), \tilde{\phi}(s) \rangle = \langle \phi(t), \phi(s) \rangle.$$

Note that the distance between the points $\phi(t)$ and $\phi(s)$ for $t, s \in T$ can be computed by using just the kernel function as;

$$\begin{aligned} d(\phi(t), \phi(s)) &= \langle \phi(t) - \phi(s), \phi(t) - \phi(s) \rangle \\ &= \langle \phi(t), \phi(t) \rangle - 2\Re\mathfrak{e}\langle \phi(t), \phi(s) \rangle + \langle \phi(s), \phi(s) \rangle \\ &= K(t,t) - 2\Re\mathfrak{e}K(t,s) + K(s,s), \quad t, s \in T. \end{aligned}$$

We would like to close this part with an illustration of how the abstract ideas developed in this presentation might be useful in handling some practical problems.

Suppose you seek a function f in a RKHS H of real valued functions on a set T with smallest norm satisfying $f(t_i) = a_i$, $i = 1, \cdots, n$, for some points $t_1, \cdots, t_n \in T$ and $a_1, \cdots, a_n \in \mathbb{R}$ (outcomes of some experiment?).
If it is not a priori clear that such a function exits, lets say you might be content to find a function in H that comes "close" to taking the given values at the specified points.

To put things in mathematical perspective, define:

$$T : H \to \mathbb{R}^n \quad \text{via} \quad T(f) := (f(t_1), \cdots, f(t_n))$$

and transform the problem to the question:

Question

Find a function $f_0 \in H$ with smallest norm that satisfies:

$$\| T(f_0) - \vec{a} \|^2 = \inf_{f \in H} \| T(f) - \vec{a} \|^2, \quad \vec{a} = (a_1, \cdots, a_n) \in \mathbb{R}^n,$$

where $\| \, . \, \|$ is the usual norm on $\mathbb{R}^n$.

Certainly, $T : H \to \mathbb{R}^n$ is linear and continuous since point evaluations on H are continuous. $T(H) = \Sigma$ is a subspace of $\mathbb{R}^n$, in particular, it is closed in $\mathbb{R}^n$. Hence there is a unique point in Σ that is closest to the point $\vec{a}$ (that is the element of smallest norm in the closed convex subset $\vec{a} + \Sigma$). This point is $P(\vec{a})$, where P is the projection onto Σ.

We would like to close this part with an illustration of how the abstract ideas developed in this presentation might be useful in handling some practical problems.

Suppose you seek a function f in a RKHS H of real valued functions on a set T with smallest norm satisfying $f(t_i) = a_i$, $i = 1, \cdots, n$, for some points $t_1, \cdots, t_n \in T$ and $a_1, \cdots, a_n \in \mathbb{R}$ (outcomes of some experiment?).
If it is not a priori clear that such a function exits, lets say you might be content to find a function in H that comes "close" to taking the given values at the specified points.

To put things in mathematical perspective, define:

$$T : H \to \mathbb{R}^n \quad \text{via} \quad T(f) := (f(t_1), \cdots, f(t_n))$$

and transform the problem to the question:

Question

Find a function $f_0 \in H$ with smallest norm that satisfies:

$$\| T(f_0) - \vec{a} \|^2 = \inf_{f \in H} \| T(f) - \vec{a} \|^2, \quad \vec{a} = (a_1, \cdots, a_n) \in \mathbb{R}^n,$$

where $\| \, . \, \|$ is the usual norm on $\mathbb{R}^n$.

Certainly, $T : H \to \mathbb{R}^n$ is linear and continuous since point evaluations on H are continuous. $T(H) = \Sigma$ is a subspace of $\mathbb{R}^n$, in particular, it is closed in $\mathbb{R}^n$. Hence there is a unique point in Σ that is closest to the point $\vec{a}$ (that is the element of smallest norm in the closed convex subset $\vec{a} + \Sigma$). This point is $P(\vec{a})$, where P is the projection onto Σ.

However, there may be many elements of H that are mapped to this point; in fact, if f is such an element, all the others form the set $f + Ker(T)$. Since this set is a closed (due to T being continuous), and also a convex set in H; it has a unique point with the least norm. Hence our problem has a unique solution. This solution u is in $(Ker(T))^{\perp}$ by the general theory, otherwise the decomposition of u into $KerT$ and $(KerT)^{\perp}$ produces an element of $f + KerT$ that has norm less than u. Now:

$$T(f) = (f(t_1), \cdots, f(t_n)) = (\langle f, \mathbb{k}_{t_1} \rangle, \cdots, \langle f, \mathbb{k}_{t_n} \rangle),$$

so for a $\vec{\zeta} = (\zeta_1, \cdots, \zeta_n)$ and $f \in H$, we compute using the inner product in $\mathbb{R}^n$;

$$\begin{aligned}\langle \vec{\zeta}, T(f) \rangle &= \sum_{i=1}^{n} \zeta_i \langle f, \mathbb{k}_{t_i} \rangle \\ &= \left\langle f, \sum_{i=1}^{n} \zeta_i \mathbb{k}_{t_i} \right\rangle \quad \forall f \in H \\ &= \langle T^*(\vec{\zeta}), f \rangle,\end{aligned}$$

where the last two inner products are in H and as usual $\mathbb{k}_{t_i}(\cdot) = K(\cdot, t_i) \in H$.

Hence we get the formula:

$$T^*(\vec{\zeta}) = \sum_{i=1}^{n} \zeta_i \mathbb{k}_{t_i}$$

However, there may be many elements of H that are mapped to this point; in fact, if f is such an element, all the others form the set $f + Ker(T)$. Since this set is a closed (due to T being continuous), and also a convex set in H; it has a unique point with the least norm. Hence our problem has a unique solution. This solution u is in $(Ker(T))^{\perp}$ by the general theory, otherwise the decomposition of u into $KerT$ and $(KerT)^{\perp}$ produces an element of $f + KerT$ that has norm less than u. Now:

$$T(f) = (f(t_1), \cdots, f(t_n)) = (\langle f, \mathbb{k}_{t_1} \rangle, \cdots, \langle f, \mathbb{k}_{t_n} \rangle),$$

so for a $\vec{\zeta} = (\zeta_1, \cdots, \zeta_n)$ and $f \in H$, we compute using the inner product in $\mathbb{R}^n$;

$$\begin{aligned} \langle \vec{\zeta}, T(f) \rangle &= \sum_{i=1}^{n} \zeta_i \langle f, \mathbb{k}_{t_i} \rangle \\ &= \left\langle f, \sum_{i=1}^{n} \zeta_i \mathbb{k}_{t_i} \right\rangle \quad \forall f \in H \\ &= \langle T^*(\vec{\zeta}), f \rangle, \end{aligned}$$

where the last two inner products are in H and as usual $\mathbb{k}_{t_i}(\cdot) = K(\cdot, t_i) \in H$.

Hence we get the formula:

$$\boxed{T^*(\vec{\zeta}) = \sum_{i=1}^{n} \zeta_i \mathbb{k}_{t_i}}$$

Since we are looking at the solution of the equation:

$$Tu = P(\vec{a}); \quad u \in (KerT)^{\perp} = R(T^*),$$

the above observation reduces our problem to finite dimensional linear algebra since $R(T^*)$ is finite dimensional and is spanned by $\mathbb{k}_{t_1}, \cdots, \mathbb{k}_{t_n}$.

In other words, our problem reduces to finding $c_1, \cdots, c_n$ of real numbers such that:

$$P(\vec{a}) = T\Big(\sum_{i=1}^{n} c_i \mathbb{k}_{t_i}\Big) = \sum_{i=1}^{n} c_i T(\mathbb{k}_{t_i}) = \sum_{i=1}^{n} c_i (\langle \mathbb{k}_{t_i}, \mathbb{k}_{t_1} \rangle, \cdots, \langle \mathbb{k}_{t_i}, \mathbb{k}_{t_n} \rangle) \quad (5)$$

$$\Big(\sum_{i=1}^{n} c_i \langle \mathbb{k}_{t_i}, \mathbb{k}_{t_1} \rangle, \cdots, \sum_{i=1}^{n} c_i \langle \mathbb{k}_{t_i}, \mathbb{k}_{t_n} \rangle\Big) = A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

$$\text{with} \qquad A = \{K(t_i, t_j)\}_{i,j=1}^{n},$$

where K is the kernel of the Hilbert space H.

However, the right hand side of the equation involves the projection of $\vec{a}$ onto Σ, which is not readily computable.

To get around this, apply T^* to both sides of Equation (5) to get:

$$\sum_{i=1}^{n} a_i \mathbb{k}_{t_i} = T^*(\vec{a}) = T^* P(\vec{a}) = \sum_{i=1}^{n} (A\vec{c})_i \mathbb{k}_{t_i}.$$

Since we are looking at the solution of the equation:

$$Tu = P(\vec{a}); \quad u \in (KerT)^{\perp} = R(T^*),$$

the above observation reduces our problem to finite dimensional linear algebra since $R(T^*)$ is finite dimensional and is spanned by $\mathbb{k}_{t_1}, \cdots, \mathbb{k}_{t_n}$.

In other words, our problem reduces to finding $c_1, \cdots, c_n$ of real numbers such that:

$$P(\vec{a}) = T\Big(\sum_{i=1}^{n} c_i \mathbb{k}_{t_i}\Big) = \sum_{i=1}^{n} c_i T(\mathbb{k}_{t_i}) = \sum_{i=1}^{n} c_i(\langle \mathbb{k}_{t_i}, \mathbb{k}_{t_1}\rangle, \cdots, \langle \mathbb{k}_{t_i}, \mathbb{k}_{t_n}\rangle) \tag{5}$$

$$\Big(\sum_{i=1}^{n} c_i\langle \mathbb{k}_{t_i}, \mathbb{k}_{t_1}\rangle, \cdots, \sum_{i=1}^{n} c_i\langle \mathbb{k}_{t_i}, \mathbb{k}_{t_n}\rangle\Big) = A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

$$\text{with} \qquad A = \{K(t_i, t_j)\}_{i,j=1}^{n},$$

where K is the kernel of the Hilbert space H.

However, the right hand side of the equation involves the projection of $\vec{a}$ onto Σ, which is not readily computable.

To get around this, apply T^* to both sides of Equation (5) to get:

$$\sum_{i=1}^{n} a_i \mathbb{k}_{t_i} = T^*(\vec{a}) = T^*P(\vec{a}) = \sum_{i=1}^{n} (A\vec{c})_i \mathbb{k}_{t_i}.$$

So solution to the equation

$$A(\vec{c}) = \vec{a}$$

will be the solution to our problem.

In the case $A = \{K(t_i, t_j)\}$ is invertible, one immediately computes the solution.

Note that A is invertible in case the positive function K is **positive definite**; that is, for every $(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k$ and $x_1, \cdots, x_k \in H$,

$$\sum_{i,j=1}^{k} \lambda_i \lambda_j K(x_i, x_j) \geq 0 \quad \text{and} \quad \sum_{i,j=1}^{k} \lambda_i \lambda_j K(x_i, x_j) = 0 \Longleftrightarrow (\lambda_1, \cdots, \lambda_k) = \vec{0}.$$

Note that the solution to this problem in case the kernel is positive definite involves only the kernel function of the Hilbert space.

Concluding Remarks

In the course of this presentation we have associated to a reproducing kernel Hilbert space H of functions on T, a positive function $K : T \times T \to \mathbb{R}$, which we called the **kernel** of H, and observed that some of the problems involving H can be solved by the use of just the kernel and nothing else.

The coming lectures will make this statement more precise. You will see that a positive function on $K : T \times T \to \mathbb{C}$ for a set K defines a reproducing kernel Hilbert space of functions on T whose kernel is precisely K.

So solution to the equation

$$A(\vec{c}) = \vec{a}$$

will be the solution to our problem.

In the case $A = \{K(t_i, t_j)\}$ is invertible, one immediately computes the solution.

Note that A is invertible in case the positive function K is **positive definite**; that is, for every $(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k$ and $x_1, \cdots, x_k \in H$,

$$\sum_{i,j=1}^{k} \lambda_i \lambda_j K(x_i, x_j) \geq 0 \quad \text{and} \quad \sum_{i,j=1}^{k} \lambda_i \lambda_j K(x_i, x_j) = 0 \iff (\lambda_1, \cdots, \lambda_k) = \vec{0}.$$

Note that the solution to this problem in case the kernel is positive definite involves only the kernel function of the Hilbert space.

Concluding Remarks

In the course of this presentation we have associated to a reproducing kernel Hilbert space H of functions on T, a positive function $K : T \times T \to \mathbb{R}$, which we called the **kernel** of H, and observed that some of the problems involving H can be solved by the use of just the kernel and nothing else.

The coming lectures will make this statement more precise. You will see that a positive function on $K : T \times T \to \mathbb{C}$ for a set K defines a reproducing kernel Hilbert space of functions on T whose kernel is precisely K.

"The Pillars of Infinite Dimensional Linear Algebra"

(In the Context of Hilbert Spaces)

(I) Uniform Boundedness Principle

Let H be a Hilbert space. Given a collection of elements $\{x_\alpha\}_{\alpha \in T}$ in H with the following property:

$$\forall x \in H \, \exists C = C(x) < \infty \text{ s.t. } \sup_{\alpha \in T} | \langle x_\alpha, x \rangle | \leq C,$$

then $\exists C > 0$ such that $\| x_\alpha \| \leq C \ \forall x_\alpha \in T$.

(II) Closed Graph Theorem

Let $T : H_1 \to H_2$ be a linear map between two Hilbert spaces. If the graph of T is closed, i.e., $x_n \to x$ and $T(x_n) \to y$ for a sequence $\{x_n\}$ in H_1 with $x \in H_1$ and $y \in H_2$ implies $T(x) = y$, then T is continuous.

(III) Alaoglu Theorem

Given a bounded sequence $\{x_n\}$ in a Hilbert space, i.e., $\exists C > 0$ such that $\| x_n \| \leq C \ \forall n$, one can find a subsequence $\{x_{k_n}\}$ of $\{x_n\}$ and a point $x \in H$ such that $\langle x_{k_n}, h \rangle \to \langle x, h \rangle$ for every $h \in H$.

"The Pillars of Infinite Dimensional Linear Algebra"
(In the Context of Hilbert Spaces)

(I) Uniform Boundedness Principle

Let H be a Hilbert space. Given a collection of elements $\{x_\alpha\}_{\alpha \in T}$ in H with the following property:

$$\forall x \in H \, \exists C = C(x) < \infty \text{ s.t. } \sup_{\alpha \in T} | \langle x_\alpha, x \rangle | \leq C,$$

then $\exists C > 0$ such that $\| x_\alpha \| \leq C \; \forall x_\alpha \in T$.

(II) Closed Graph Theorem

Let $T : H_1 \to H_2$ be a linear map between two Hilbert spaces. If the graph of T is closed, i.e., $x_n \to x$ and $T(x_n) \to y$ for a sequence $\{x_n\}$ in H_1 with $x \in H_1$ and $y \in H_2$ implies $T(x) = y$, then T is continuous.

(III) Alaoglu Theorem

Given a bounded sequence $\{x_n\}$ in a Hilbert space, i.e., $\exists C > 0$ such that $\| x_n \| \leq C \; \forall n$, one can find a subsequence $\{x_{k_n}\}$ of $\{x_n\}$ and a point $x \in H$ such that $\langle x_{k_n}, h \rangle \to \langle x, h \rangle$ for every $h \in H$.

"The Pillars of Infinite Dimensional Linear Algebra"
(In the Context of Hilbert Spaces)

(I) Uniform Boundedness Principle

Let H be a Hilbert space. Given a collection of elements $\{x_\alpha\}_{\alpha \in T}$ in H with the following property:

$$\forall x \in H \, \exists C = C(x) < \infty \text{ s.t. } \sup_{\alpha \in T} | \langle x_\alpha, x \rangle | \leq C,$$

then $\exists C > 0$ such that $\| x_\alpha \| \leq C \; \forall x_\alpha \in T$.

(II) Closed Graph Theorem

Let $T : H_1 \to H_2$ be a linear map between two Hilbert spaces. If the graph of T is closed, i.e., $x_n \to x$ and $T(x_n) \to y$ for a sequence $\{x_n\}$ in H_1 with $x \in H_1$ and $y \in H_2$ implies $T(x) = y$, then T is continuous.

(III) Alaoglu Theorem

Given a bounded sequence $\{x_n\}$ in a Hilbert space, i.e., $\exists C > 0$ such that $\| x_n \| \leq C \; \forall n$, one can find a subsequence $\{x_{k_n}\}$ of $\{x_n\}$ and a point $x \in H$ such that $\langle x_{k_n}, h \rangle \to \langle x, h \rangle$ for every $h \in H$.

(IV) Spectral Theorem for Compact Operators

Given a linear, continuous operator T from a separable Hilbert space H into itself with the additional properties:

- **(*)** For every bounded sequence $\{x_n\}$ in H, there exists a subsequence $\{T(x_{k_n})\}$ of $\{T(x_n)\}$ such that $\{T(x_{k_n})\}$ converges in H,
- **(**)** $T^* = T$,

then there exists a sequence λ_n of real numbers converging to zero and an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ of H such that

$$T(x) = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \quad \text{in } H.$$

Acknowledgment

My sincere thanks goes to Buket Can Bahdır, for creating this beamer presentation out of my handwritten notes. Thank you Buket.

- H. Aronszajn, Theory of Reproducing Kernels, TAMS Vol. 68, No.3, 1950, pp. 337-404.
- E. Kreyszig, Introductory Functional Analysis with Applications, John Wiley & Sons, 1978.
- R. Meise, D. Vogt, Introduction to Functional Analysis, Oxford Graduate Texts in Mathematics, Band 2, 1997.
- J.H. Manton, P. Amblard, A Primer on Reproducing Kernel Hilbert Spaces, Foundations and Trends in Signal Processing Vol. XX, No XX, 2015.
- D. Okutmuştur, A Survey on Hilbert Spaces and Reproducing Kernels, In Functional Calculus, London: Intech Open, 2020, pp 61-77.
- V. Paulsen, M. Raghupathi, An Introduction to the Theory of Reproducing Kernel Hilbert Spaces, Cambridge University Press 2016.
- W. Rudin, Functional Analysis, 2. Edition, Mc Graw-Hill.