11.0 Hilbert Spaces

11.1 Definitions of Properties

Definition: An *inner product space* or a *pre-Hilbert space* is a linear space H with an inner product operator

$$< \cdot >: H \times H \to \mathbf{C}$$

that maps

$$(x, y) \to < x, y >$$

that satisfies the following properties:

1. $< x, y > = < y, x >^*$.
2. $< ax_1 + bx_2, y > = a < x_1, y > +b < x_2, y > \quad \forall\ a, b \in \mathbf{C}$.
3. $< x, x > \geq 0\ \forall\ x \in H$ and $< x, x > = 0$ if and only if $x = 0$ (0 in H).

Definition: The *norm* of $x \in H$ is defined to be

$$||x|| := \sqrt{< x, x >}.$$

Example: Let $H = L_2(a, b)$, the set of all functions defined on (a, b) satisfying

$$\int_a^b |f(x)|^2 dx < \infty$$

with

$$< f, g > := \int_a^b f(x) g^*(x) dx$$

and

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

Then L_2 is a pre-Hilbert space.

Definition: Let H be a pre-Hilbert space. A sequence of elements from H denoted, $x_1, x_2, \ldots$, is said to be a *Cauchy sequence* if

$$\lim_{n,m \to \infty} ||x_n - x_m|| = 0.$$

Definition: A pre-Hilbert space H is said to be *complete* if every Cauchy sequence converges to an element in H, i.e., there exists $x \in H$ such that

$$\lim_{n\to\infty} ||x_n - x|| = 0.$$

Definition: A *Hilbert space* is a complete pre-Hilbert space.

Note: The pre-Hilbert space L_2 defined earlier is a Hilbert space.

Example: Let S be the set of all random variables defined on some probability space (U, F, P) such that

$$E\left[|X(u)|^2\right] < \infty \ \forall \ x \in S.$$

Let us say $X \sim Y$ if and only if

$$E\left[|X(u) - Y(u)|^2\right] = 0.$$

Then $\sim$ is an equivalence relation since

1. $E\left[|X(u) - X(u)|^2\right] = 0$ (so $X \sim X$).
2. If $E\left[|X(u) - Y(u)|^2\right] = 0$ then $E\left[|Y(u) - X(u)|^2\right] = 0$ (so $X \sim Y \Rightarrow Y \sim X$).
3. If $E\left[|X(u) - Y(u)|^2\right] = 0$ and $E\left[|Y(u) - Z(u)|^2\right] = 0$ then

$$E\left[|X(u) - Z(u)|^2\right] = E\left[|X(u) - Y(u) + Y(u) - Z(u)|^2\right]$$

$$\leq E\left[(|X(u) - Y(u)| + |Y(u) - Z(u)|)^2\right]$$

$$= E\left[|X(u) - Y(u)|^2 + 2|X(u) - Y(u)||Y(u) - Z(u)| + |Y(u) - Z(u)|^2\right]$$

$$= 2E\left[|X(u) - Y(u)||Y(u) - Z(u)|\right].$$

In order to see this last expression is 0 consider the following (here $X = X(u)$, etc.).

Let

$$U = |X - Y|, \quad V = |Y - Z|.$$

Let

$$W = aU + V.$$

Then

$$E(W^2) = E(a^2U^2 + 2aUV + V^2) \geq 0 \;\forall a$$

but

$$E(W^2) = E(V^2) = 0$$

so

$$aE(UV) \geq 0 \;\forall a$$

thus

$$E(UV) = 0$$

(so $X \sim Y$ and $Y \sim Z \Rightarrow X \sim Z$).

Hence, $\sim$ is an equivalence relation. This equivalence relation partitions S into distinct equivalence classes. Two members of the same equivalence class are said to be mean square equivalent. What this means then is on this probability space we do not distinguish between elements X and Y if

$$E\left[|X(u) - Y(u)|^2\right] = 0.$$

Now define

$$< X, Y >:= E(XY^*)$$

and hence

$$||X|| = \sqrt{E|X^2|}.$$

We have

1. $E(XY^*) = [E(YX^*)]^*$ (so $< X, Y >=< Y, X >^*$).

2. $E[(aX_1 + bX_2)Y^*] = aE(X_1Y^*) + bE(X_2Y^*)$ (so $< aX_1 + bX_2, Y >= a < X_1, Y > +b < X_2, Y >$).

3. $E(XX^*) = E(|X|^2) \geq 0$ and $E(|X^2|) = 0$ if and only if $X = 0$ (so $< X, X >\geq 0$ and $< X, X >= 0$ if and only if $X = 0$).

Lemma (*Cauchy-Schwarz Inequality*): For all x, y in an inner product space

$$| < x, y > | \leq ||x|| \; ||y||.$$

Equality holds if and only if $x = \lambda y$ or $y = 0$, where $\lambda \in \mathbf{C}$.

Proof: If $y = 0$ then it is trivial. Suppose $y \neq 0$. For all scalars λ we have

$$0 \leq < x - \lambda y, x - \lambda y >$$

$$= < x, x > - \lambda < y, x > - \lambda^* < x, y > + |\lambda|^2 < y, y > .$$

Let

$$\lambda = \frac{< x, y >}{< y, y >}.$$

Then

$$0 \leq < x, x > - \frac{< x, y > < x, y >^*}{< y, y >}$$

$$- \frac{< x, y >^* < x, y >}{< y, y >} + \frac{< x, y >}{< y, y >} \frac{< x, y >^*}{< y, y >} < y, y >$$

$$= < x, x > - \frac{| < x, y > |^2}{< y, y >}$$

so

$$< x, y > \leq \sqrt{< x, x > < y, y >} = ||x|| \; ||y||.$$

Proposition: On a pre-Hilbert space H the function

$$||x|| = \sqrt{< x, \; x >}$$

is a valid norm, i.e.,

1. $||x|| \geq 0 \; \forall \; x \in H, \; ||x|| = 0$ if and only if $x = 0$.
2. $||x + y|| \leq ||x|| + ||y|| \; \forall \; x, y \in H$.
3. $||\alpha x|| = |\alpha| \cdot ||x|| \; \forall \; x \in H, \;$ scalar α.

Proof: Homework exercise.

11.2 The Projection Theorem

Definition: In a pre-Hilbert space two vectors x, y are said to be *orthogonal* if $< x,\ y >= 0$. Here write $x \perp y$. A vector x is said to be orthogonal to a set S ($x \perp S$) if $x \perp s$ for all $s \in S$.

Lemma: If $x \perp y$ then $||x+y||^2 = ||x||^2 + ||y||^2$.

Proof: Homework exercise.

Lemma (*Parallelogram Law*): In a pre-Hilbert space

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

Proof: Homework exercise.

Proposition: Let $x_n, y_n \in H$, a Hilbert space. If

$$\lim_{n\to\infty} ||x_n - x|| = 0 \quad \text{and} \quad \lim_{n\to\infty} ||y_n - y|| = 0$$

then

$$\lim_{n\to\infty} < x_n,\ y_n >=< x,\ y > .$$

Proof: Here $||x|| \le M$ and $||y|| \le M,\ M < \infty$.

$$< x_n,\ y_n >=< x_n - x + x,\ y_n - y + y >$$

$$=< x_n - x,\ y_n - y > + < x_n - x,\ y > + < x_n,\ y_n - y > + < x,\ y >$$

so

$$| < x_n,\ y_n > - < x,\ y > | \le | < x_n - x,\ y_n - y > | + | < x_n - x,\ y > | + | < x_n,\ y_n - y > |$$

$$\le ||x_n - x|| \cdot ||y_n - y|| + ||x_n - x|| \cdot ||y|| + ||x|| \cdot ||y_n - y||$$

and the right hand side goes to 0 as $n \to \infty$.

This is called *continuity of the norm.*

Theorem (*Classical Projection Theorem*): Let H be a Hilbert space and M a closed subspace of H. For any vector $x \in H$ there exists a unique vector $m_0 \in M$ such that

$$||x - m_0|| \leq ||x - m|| \ \forall m \in M.$$

Furthermore, a necessary and sufficient condition that $m_0 \in M$ be the unique minimizing vector is that $x - m_0$ is orthogonal to M.

Proof: First we show that if m_0 is a minimizing vector then $(x - m_0) \perp M$. Suppose there exists $m \in M$ which is not orthogonal to $x - m_0$. Without loss of generality we may assume $||m|| = 1$ and

$$< x - m_0, m >= \delta \neq 0.$$

Define $m_1 \in M$ by $m_1 = m_0 + \delta m$. Then

$$||x - m_1||^2 = ||x - m_0 - \delta m||^2$$

$$= ||x - m_0||^2 - < x - m_0, \delta m > - < \delta m, x - m_0 > + |\delta|^2 ||m||^2$$
$$= ||x - m_0||^2 - |\delta|^2 < ||x - m_0||^2.$$

Thus, if $x - m_0$ is not orthogonal to M then m_0 is not a minimizing vector. So if m_0 exists then $(x - m_0) \perp M$.

Note: In the above we used operations like

$$< x, ay >=< ay, x >^* = (a < y, x >)^* = a^* < y, x >^* = a^* < x, y >$$

so

$$< x - m_0, \delta m >= \delta^* < x - m_0, m >= \delta^* \delta = |\delta|^2$$

and

$$< \delta m, x - m_0 >= \delta < x - m_0, m >^* = \delta \delta^* = |\delta|^2.$$

We now show that if $(x - m_0) \perp M$ then m_0 is unique.

For any $m \in M$

$$||x - m||^2 = ||x - m_0 + m_0 - m||^2 = ||x - m_0||^2 + ||m_0 - m||^2.$$

Thus

$$||x - m|| > ||x - m_0|| \text{ for } m \neq m_0.$$

We now establish the existence of m_0.

If $x \in M$ then we take $m_0 = x$ and we are done. So assume now that $x \notin M$ and define

$$\Delta = \inf_{m \in M} ||x - m||.$$

We wish to produce an $m_0 \in M$ such that $||x - m_0|| = \Delta$. Let $\{m_i\}$ be a sequence of vectors in M such that $||x - m_i|| \to \Delta$. By the parallelogram law

$$||(m_j - x) + (x - m_i)||^2 + ||(m_j - x) - (x - m_i)||^2$$
$$= 2||m_j - x||^2 + 2||x - m_i||^2$$

so

$$||m_j - m_i||^2 = 2||m_j - x||^2 + 2||x - m_i||^2 - 4\left|\left|x - \frac{m_i + m_j}{2}\right|\right|^2.$$

Now

$$\frac{m_i + m_j}{2} \in M$$

since M is a closed linear space. So

$$\left|\left|x - \frac{m_i + m_j}{2}\right|\right| \geq \Delta.$$

Thus

$$||m_j - m_i||^2 \leq 2||m_j - x||^2 + 2||x - m_i||^2 - 4\Delta^2.$$

Since

$$\lim_{i \to \infty} ||m_i - x||^2 \to \Delta^2$$

we get

$$||m_j - m_i||^2 \to 0 \text{ as } i, j \to \infty.$$

Hence $\{m_i\}$ is a Cauchy sequence and since M is a closed subspace of a complete space, the sequence $\{m_i\}$ has a limit $m_0 \in M$. By continuity of the norm, $||x - m_0|| = \Delta$.

11.3 Mean-Square Convergence

Theorem: If $X_n \to X$ (m.s.) and $Y_n \to Y$ (m.s.) then

1. $\lim_{n\to\infty} E[X_n] = E[X]$.
2. $\lim_{n\to\infty} E[X_n Y_n] = E[XY]$.

Proof:

1. $$|E[X_n - X]|^2 \le E\left[|X_n - X|^2\right] \to 0 \text{ as } n \to \infty$$

   by definition of mean-square convergence so

   $$\lim_{n\to\infty} E[X_n - X] = 0 \Rightarrow \lim_{n\to\infty} E[X_n] = E[X].$$

2. $$\begin{aligned} |E[X_nY_n - XY]| &= |E[X_n(Y_n - Y) - (X - X_n)Y]| \\ &\le |E[X_n(Y_n - Y)]| + |E[(X - X_n)Y]| \\ &\le \sqrt{E[|X_n|]}\sqrt{E[|Y_n - Y|]} + \sqrt{E[|X - X_n|]}\sqrt{E[|Y|]} \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Theorem: A sequence of random variables X_n is convergent in the mean-square sense *if and only if*

$$\lim_{n,m\to\infty} R_X(n,m)$$

exists, i.e.,

$$\lim_{n,m\to\infty} E[X_n X_m^*] \text{ exists.}$$

Proof: " $\Leftarrow$ " Say $R = \lim_{n,m\to\infty} R_X(n,m)$. Then

$$E[|X_n - X_m|^2] = E[|X_n|^2] - E[|X_n X_m^*|] - E[|X_n^* X_m|] + E[|X_m|^2]$$

$$\longrightarrow R - R - R + R = 0 \Rightarrow E[|X_n - X_m|^2] \to 0$$

$$\Rightarrow E[|X_n - X_m|] \to 0 \to X_n \text{ is Cauchy} \to X_n \text{ is m.s. convergent.}$$

" $\Rightarrow$ " This follows from properties above.