Basic Facts About Hilbert Space

The term Euclidean space refers to a finite dimensional linear space with an inner product. A Euclidean space is always complete by virtue of the fact that it is finite dimensional (and we are taking the scalars here to be the reals which have been constructed to be complete). An infinite dimensional inner product space which is complete for the norm induced by the inner product is called a **Hilbert space**. A Hilbert space is in many ways like a Euclidean space (which is why finite dimensional intuituition often works in the infinite dimensional Hilbert space setting) but there are some ways in which the infinite dimensionality leads to subtle differences we need to be aware of.

Subspaces

A subset M of Hilbert space H is a **subspace** of it is closed under the operation of forming linear combinations; i.e., for all x and y in M, $C_1x + C_2y$ belongs to M for all scalars C_1, C_2. The subspace M is said to be **closed** if it contains all its limit points; i.e., every sequence of elements of M that is Cauchy for the H-norm, converges to an element of M. In a Euclidean space every subspace is closed but in a Hilbert space this is not the case.

Examples-

(a) If U is a bounded open set in R^n then $H = H^0(U)$ is a Hilbert space containing $M = C(U)$ as a subspace. It is easy to find a sequence of functions in M that is Cauchy for the H norm but the sequence converges to a function in H that is discontinuous and hence not in M. This proves that M is not closed in H.

(b) Every finite dimensional subspace of a Hilbert space H is closed. For example, if M denotes the span of finitely many elements $x_1, \ldots .x_N$ in H, then the set M of all possible linear combinations of these elements is finite dimensional (of dimension N), hence it is closed in H.

(c) Let M denote a subspace of Hilbert space H and let $M^\perp$ denote the **orthogonal complement of M**.

$$M^\perp = \{x \in H : (x,y)_H = 0, \forall y \in M\}$$

Then $M^\perp$ is easily seen to be a subspace and it is closed, whether or not M itself is closed. To see this, suppose $\{x_n\}$ is a Cauchy sequence in $M^\perp$ converging to limit x in H. For arbitrary y in M,

$$(x,y)_H = (x - x_n, y)_H + (x_n, y)_H = (x - x_n, y)_H + 0 \to 0, \text{ as n tends to infinity.}$$

Then the limit point x is orthogonal to every y in M which is to say, x is in $M^\perp$, and $M^\perp$ is closed.

Lemma 1- Let M denote a subspace of Hilbert space H. Then $(M^\perp)^\perp = \bar{M}$.

If M is a subspace of H that is not closed, then M is contained in a closed subspace $\bar{M}$ of H, consisting of M together with all its limit points. $\bar{M}$ is called the **closure** of M and M is said to be **dense** in $\bar{M}$. This means that for every x in $\bar{M}$ there is a sequence of elements of M that converge to x in the norm of H. Equivalently, to say M is dense in $\bar{M}$ means that for every x in $\bar{M}$ and every $\epsilon > 0$, there is a y in M such that $\|x-y\|_H < \epsilon$.

Lemma 2 A subspace M of Hilbert space H is dense in H if and only if $M^{\perp} = \{0\}$.

A Hilbert space H is said to be **separable** if H contains a countable dense subset $\{h_n\}$. In this case, for every x in H and every $\epsilon > 0$ there exists an integer N_ϵ and scalars $\{a_n\}$ such that

$$\left\| x - \sum_{n=1}^{N} a_n h_n \right\|_H < \epsilon \quad \text{for} \quad N > N_\epsilon$$

If H is a separable Hilbert space, then the Gram-Schmidt procedure can be used to construct an orthonormal basis for H out of a countable dense subset. An orthonormal basis for H is a set of mutually orthogonal unit vectors, $\{\phi_n\}$ in H with the following property:

1) For $f \in H$, $(\phi_n, f)_H = 0$ for every n if and only if $f = 0$

When the orthonormal set $\{\phi_n\}$ has property 1, then it is said to be **dense** or **complete** in H. Of course, not every orthonormal set in H is complete. Other equivalent ways of characterizing completeness for orthonormal sets can be stated as follows:

2) For all f in H and every $\epsilon > 0$, there exists an integer N_ϵ such that

$$\left\| f - \sum_{n=1}^{N} (f,\phi_n)_H \phi_n \right\|_H < \epsilon \quad \text{for} \quad N > N_\epsilon$$

3) For every f in H, $\sum_{n=1}^{\infty} f_n^2 = \|f\|_H^2$ where $f_n = (f,\phi_n)_H$

In a Euclidean space, E, where all subspaces M are closed, it is a fact that for each y in E there is a unique z in M such that $\|y-z\|$ is minimal. This element z, which is just the orthogonal projection of y onto M, is the ”best approximation to y from within M”. In an infinite dimensional Hilbert space, a similar result is true for closed subspaces but for subspaces that are not closed there may fail to be a ”best” approximation in M.

Hilbert Space Projection Theorem Let M be a closed subspace of Hilbert space H and let y in H be given. Then

(i) there exists a unique x_y in M such that $\|y-x_y\|_H \leq \|y-z\|_H$ for all z in M
(x_y is the unique point of M that is closest to y, the best approximation in M to y)

(ii) $(y-x_y,z)_H = 0$ for all z in M; i.e., $y-x_y \perp M$

(iii) every y in H can be uniquely expressed as $y = x_y + z_y$ where

$$Py = x_y \in M,\ \ Qy = z_y \in M^{\perp}$$

and $\|y\|_H^2 = \|Py\|_H^2 + \|Qy\|_H^2$ i.e., $H = M \oplus M^{\perp}$.

The proof of this result will be given later.

Linear Functionals and Bilinear Forms

A real valued function defined on H, is said to be a ***functional*** on H. The functional, L, is said to be:

(a) **Linear** if, for all x and y in H, $L(C_1x + C_2y) = C_1Lx + C_2Ly$, for all scalars C_1, C_2.

(b) **Bounded** if there exists a constant C such that $|Lx| \leq C\|x\|_H$ for all x in H

(c) **Continuous** if $\|x_n - x\|_H \to 0$ implies that $|Lx_n - Lx| \to 0$

It is not difficult to show that the only example of a linear functional on a Euclidean space E is $Lx = (x,z)_E$ for some z in E, fixed. For example, if F is a linear functional on E, then for arbitrary x in E,

$$F(x) = F\left(\sum_{i=1}^n x_i\bar{e}_i\right) = \sum_{i=1}^n x_iF(\bar{e}_i) = \sum_{i=1}^n x_iF_i = (x,z_F)_E = x^{\top}z_F$$

where $\{e_i\}$ denotes the standard basis in E and $\bar{z}_F$ denotes the n-tuple whose i-th component is $F_i = F(\bar{e}_i)$. This displays the isomorphism between functionals F and elements, z_F, in E. This isomorphism also exists in an abstract Hilbert space.

Riesz Representation Theorem For every continuous linear functional f on Hilbert space H there exists a unique element z_f in H such that $f(x) = (x,z_f)_H$ for all x in H.

Proof- Let $N_f = \{x \in H : f(x) = 0\}$.Then N_f is easily seen to be a closed subspace of H. If $N_f = H$ then $z_f = 0$ and we are done. If $N_f \neq H$ then $H = N_f \oplus N_f^{\perp}$ by the Hilbert space projection theorem. Since N_f is not all of H, $N_f^{\perp}$ must contain nonzero vectors, and we denote by z_0 an element of $N_f^{\perp}$ such that $\|z_0\|_H = 1$. Then for any x in H,

$$w = f(x)z_0 - f(z_0)x$$

belongs to N_f hence $w \perp z_0$. But in that case,

$$(f(x)z_0 - f(z_0)x, z_0)_H = f(x)(z_0,z_0)_H - f(z_0)(x,z_0)_H = 0.$$

This leads to, $f(x) = f(z_0)(x,z_0)_H = (x,f(z_0)z_0)_H$ which is to say $z_f = f(z_0)z_0$.

To see that z_f is unique, suppose that $f(x) = (z_f,x)_H = (w_f,x)_H$ for all x in H. Subtracting leads to the result that $(z_f - w_f,x)_H = 0$ for all x in H. In particular, choosing $x = z_f - w_f$ leads to $\|(z_f - w_f\|_H = 0.$ ■

A real valued function $a(x,y)$ defined on $H \times H$ is said to be:

(a) **Bilinear** if, for all $x_1,x_2,y_1,y_2 \in H$ and all scalars C_1, C_2

$$a(C_1x_1 + C_2x_2,y_1) = C_1a(x_1,y_1) + C_2a(x_2,y_1)$$

$$a(x_1,C_1y_1 + C_2y_2) = C_1a(x_1,y_1) + C_2a(x_2,y_1)$$

(b) **Bounded** if there exists a constant $b > 0$ such that,

$|a(x,y)| \leq b\|x\|_H\|y\|_H$ for all x,y in H

(c) **Continuous** if $x_n \to x,$ and $y_n \to y$ in H, implies $a(x_n,y_n) \to a(x,y)$ in R

(d) **Symmetric** if $a(x,y) = a(y,x)$ for all x,y in H

(e) **Positive** or **coercive** if there exists a $C > 0$ such that

$a(x,x) \geq C\|x\|_H^2$ for all x in H

It is not hard to show that for both linear functionals and bilinear forms, boundedness is equivalent to continuity. If a(x,y) is a bilinear form on $H \times H$, and F(x) is a linear functional on H, then $\Phi(x) = a(x,x)/2 - F(x) + Const$ is called a **quadratic functional on H**. In a Euclidean space a quadratic functional has a unique extreme point located at the point where the gradient of the functional vanishes. This result generalizes to the infinite dimensional situation.

Lemma 3 Suppose $a(x,y)$ is a positive, bounded and symmetric bilinear form on Hilbert space H, and F(x) is a bounded linear functional on H. Consider the following problems

(a) minimize $\Phi(x) = a(x,x)/2 - F(x) + Const$ over H

(b) find x in H satisfying $a(x,y) = F(y)$ for all y in H.

Then

i) x in H solves (a) if and only if x solves (b)

ii) there is at most on x in H solving (a) and (b)

iii) there is at least one x in H solving (a) and (b)

Proof- For t in R and x,y fixed in H, let $f(t) = \Phi(x+ty)$. Then f(t) is a real valued function of the real variable t and it follows from the symmetry of $a(x,y)$ that

$$f(t) = t^2/2\ a(y,y) + t[a(x,y) - F(y)] + 1/2\ a(x,x) - F(x) + Const$$

and

$$f'(t) = t\ a(y,y) + [a(x,y) - F(y)]$$

It follows that $\Phi(x)$ has a global minimum at x in H if and only if $f(t)$ has a global minimum at $t = 0$; i.e.,

$$\Phi(x+ty) = \Phi(x) + tf'(0) + t^2/2\ a(x,x) \geq \Phi(x), \quad \forall t \in R \quad and \quad \forall y \in H$$

if and only if

$$f'(0) = a(x,y) - F(y) = 0. \quad \forall y \in H.$$

This establishes the equivalence of (a) and (b).

To show that $\Phi(x)$ has at most one minimum in H, suppose

$a(x_1,y) = F(y)$ and $a(x_2,y) = F(y)$ for all y in H.

Then $a(x_1,y) - a(x_2,y) = a(x_1 - x_2, y) = 0$ for all y in H. In particular, for $y = x_1 - x_2$

$0 = a(x_1 - x_2, x_1 - x_2) \geq C\|x_1 - x_2\|_H^2$; i.e., $x_1 = x_2$

To show that $\Phi(x)$ has at least one minimum in H, let $\alpha = \inf_{x \in H} \Phi(x)$. Now

$$\Phi(x) = 1/2\ a(x,x) - F(x) \geq 1/2\ C\|x\|_H^2 - b\|x\|_H$$

and it is evident that $\Phi(x)$ tends to infinity as $\|x\|_H$ tends to infinity. This means $\alpha > -\infty$ (i.e.," the parabola opens upward rather than downward"). Moreover since α is an infimum, there exists a sequence x_n in H such that $\Phi(x_n) \to \alpha$ as n tends to infinity. Note that

$$2[a(x_n,x_n) + a(x_m,x_m)] = a(x_n - x_m, x_n - x_m) + a(x_m + x_n, x_m + x_n)$$

which leads to the result,

$$\Phi(x_m) + \Phi(x_n) = 1/4\ a(x_m - x_n, x_m - x_n) + 2\ \Phi[(x_m + x_n)/2] \geq 1/4C\|x_m - x_n\|_H^2 + 2\alpha.$$

But $\Phi(x_m) + \Phi(x_n)$ tends to 2α as n tends to infinity and in view of the previous line, the minimizing sequence $\{x_n\}$ must be a Cauchy sequence with limit x in the Hilbert space H. Finally, since $\Phi(x)$ is continuous, $\Phi(x_n) \to \Phi(x) = \alpha$.■

Applications of the lemma-

(i) This lemma can now be used to prove the Hilbert space projection theorem.
For M a closed subspace in H and for y a fixed but arbitrary element in H, note that

$$\|x-y\|_H^2 = (x-y,x-y)_H = \|x\|_H^2 - 2(x,y)_H + \|y\|_H^2 \quad \text{for all x in M.}$$

Since M is closed in H, it follows that M is itself a Hilbert space for the norm and inner product inherited from H.
Define

$$a(z,x) = (z,x)_H \qquad \text{for x and z in M,}$$
$$F(z) = (y,z)_H \qquad \text{for z in M,}$$
and
$$\Phi(z) = 1/2\ a(z,z) - F(z) + 1/2\|y\|_H^2 \qquad \text{for z in M}$$

Clearly $a(z,x)$ is a positive, bounded and symmetric bilinear form on M, F is a bounded linear functional on M. Then it follows from the lemma that there exists a unique element $x_y \in M$ which minimizes $\Phi(z)$ over M. It follows also form the equivalence of problems (a) and (b) that x_y satisfies $a(x_y,z) = F(z)$, for all z in M; i.e., $(x_y,z)_H = (y,z)_H$ for all z in M. But this is just the assertion that $(x_y - y,z)_H = 0$ for all z in M, that is, $x_y - y \perp M$. Finally, for y in H, fixed, let the unique element x_y in M be denoted by $Py = x_y \in M$. Then

$$y - Py \perp M, \text{ and } z = y - Py \in M^\perp.$$

To see that this decomposition of elements of H is unique, suppose

$$y = x_y + z, \quad x_y \in M, \quad z \in M^\perp,$$
and
$$y = X_y + Z, \quad X_y \in M, \quad Z \in M^\perp,$$
Then
$$x_y + z = X_y + Z, \text{ and } x_y - X_y = Z - z.$$

But $\quad x_y - X_y \in M, \quad Z - z \in M^\perp, \quad M \cap M^\perp = \{0\},$
and it follows that $\quad x_y - X_y = Z - z = 0.$■

(ii) Recall that for U open and bounded in R^n, $H_0^1(U) = M$, is a closed subspace of $H^1(U) = H$. Then by the projection theorem, every y in H can be uniquely expressed as a sum, $y = x_y + z$, with $x_y \in M$, and $z \in M^\perp$. To characterize the subspace $M^\perp$, choose arbitrary $\phi \in C_0^\infty(U)$ and $\psi \in C^\infty(U)$ and write

$$(\phi,\psi)_H = \int_U [\phi\psi + \nabla\phi \bullet \nabla\psi]dx = \int_U \phi[\psi - \nabla^2\psi]dx + \int_{\partial U} \phi\partial_N\psi dS$$

$$= (\phi, \psi - \nabla^2\psi)_0 + 0. \text{ (Recall that } (u,v)_0 \text{ denotes the } H^0(U) \text{ inner product).}$$

Now suppose $\psi \in C^{\infty}(U) \cap M^{\perp}$. Then $(\phi,\psi)_H = 0,$ for all $\phi \in C_0^{\infty}(U),$ and since $C_0^{\infty}(U)$ is dense in M, $(u,\psi)_H = 0,$ for all u in M. That is, $(u,\psi - \nabla^2\psi)_0 = 0$ for all u in M. But this implies that $\psi \in C^{\infty}(U) \cap M^{\perp}$ satisfies $\psi - \nabla^2\psi = 0,$ in $H^0(U)$. Then, since $C^{\infty}(U)$ is dense in $H = H^1(U)$ (cf. Theorem 2 pg 250 in the text) it follows that

$$M^{\perp} = \{z \in H : z - \nabla^2 z \in H^0(U), \text{ and } z - \nabla^2 z = 0\}$$

The lemma requires that the bilinear form $a(x,y)$ be symmetric. For application to existence theorems for partial differential equations, this is an unacceptable restriction. Fortunately, the most important part of the result remains true even when the form is not symmetric.

For A an n by n, not necessarily symmetric, but positive definite matrix, consider the problem Ax = f. For any n by n matrix $dimN_A = dimN_{A^\top},$ and for A positive definite, $dimN_A = 0,$ which is to say that the solution of $Ax = f$ is unique if it exists. Since $R^n = R_A \oplus N_{A^\top},$ it follows that $R^n = R,$ which is to say, a solution for $Ax = f$ exists for every f in R^n. The situation in an abstract Hilbert space H is very close to this.

Lax-Milgram Lemma- Suppose a(u,v) is a positive and bounded bilinear form on Hilbert space H; i.e.,

$$|a(u,v)| \leq \alpha \|u\|_H \|v\|_H \quad \forall u,v \in H$$

and

$$a(u,u) \geq \beta \|u\|_H^2 \quad \forall u \in H.$$

Suppose also that $F(v)$ is a bounded linear functional on H. Then there exists a unique U in H such that

$$a(U,v) = F(v) \quad \forall v \in H.$$

Proof- For each fixed u in H, the mapping $v \rightsquigarrow a(u,v)$ is a bounded linear functional on H. It follows that there exists a unique $z_u \in H$ such that

$$a(u,v) = (z_u,v)_H \quad \forall v \in H.$$

Let $Au = z_u$; i.e., $a(u,v) = (Au,v)_H \ \ \forall u \in H.$ Clearly A is a linear mapping of H into H, and since

$$\|Au\|_H^2 = |(Au,Au)_H| = |a(u,Au)| \leq \alpha \|u\|_H \|Au\|_H$$

it is evident that A is also bounded. Note further, that

$$\beta \|u\|_H^2 \leq a(u,u) = (Au,u)_H \leq \|Au\|_H \|u\|_H$$

i.e., $\quad \beta \|u\|_H \leq \|Au\|_H \quad \forall u \in H.$

This estimate implies that A is one-to one and that R_A, the range of A, is closed in H. Finally, we will show that $R_A = H$. Since the range is closed, we can use the projection theorem to write, $H = R_A \oplus R_A^{\perp}$. If $u \in R_A^{\perp}$, then

$$0 = (Au,u)_H = a(u,u) \geq \beta\|u\|_H^2; \quad \text{i.e., } R_A^{\perp} = \{0\}.$$

Since $F(v)$ is a bounded linear functional on H, it follows from the Riesz theorem that there is a unique $z_F \in H$ such that $F(v) = (z_F,v)_H$ for all v in H. Then the equation $a(u,v) = F(v)$ can be expressed as

$$(Au,v)_H = (z_F,v)_H \quad \forall v \in H; \text{ i.e., } Au = z_F.$$

But A has been seen to be one-to-one and onto and it follows that there exists a unique $U \in H$ such that $AU = z_F$. ■

Convergence

In R^N convergence of x_n to x means

$$\|x_n - x\|_{R^N} = \left[\sum_{i=1}^{N}[(x_n - x)\bullet e_i]^2\right]^{1/2} \to 0 \text{ as } n \to \infty.$$

Here e_i denotes the i-th vector in the standard basis.This is equivalent to,

$$(x_n - x)\bullet e_i \to 0 \text{ as } n \to \infty, \text{ for } i = 1,...,N,$$

and to
$$(x_n - x)\bullet z \to 0 \text{ as } n \to \infty, \text{ for every } z \in R^N$$

In an infinite dimensional Hilbert space H, convergence of x_n to x means $\|x_n - x\|_H \to 0$ as $n \to \infty$. This is called **strong convergence** in H and it implies that

$$(x_n - x,v)_H \to 0 \text{ as } n \to \infty \quad \forall v \in H.$$

This last mode of convergence is referred to as **weak convergence** and, in a general Hilbert space, weak convergence does not imply strong convergence. Thus while there is no distinction between weak and strong convergence in a finite dimensional space, the two notions of convergence are not the same in a space of infinite dimensions.

In R^N, a sequence $\{x_n\}$ is said to be bounded if there is a constant M such that $|x_n| \leq M$ for all n. Then the Bolzano-Weierstrass theorem asserts that every bounded sequence $\{x_n\}$ contains a convergent subsequence. To see why this is true, note that $\{x_n \bullet e_1\}$ is a bounded sequence of real numbers and hence contains a subsequence $\{x_{n,1}\bullet e_1\}$ that is convergent. Similarly, $\{x_{n,1}\bullet e_2\}$ is also a bounded sequence of real numbers and thus contains a subsequence $\{x_{n,2}\bullet e_2\}$ that is convergent. Proceeding in this way, we can generate a sequence of subsequences, $\{x_{n,k}\} \subset \{x_n\}$ such that $\{x_{n,k}\bullet e_j\}$ is convergent for $j \leq k$. Then the diagonal sequence $\{x_{n,n}\}$ is such that $\{x_{n,n}\bullet ej\}$ is convergent for

$1 \le j \le N,$ which is to say, $\{x_{n,n}\}$ is convergent.

The analogue of this result in a Hilbert space is the following.

Lemma 4- Every bounded sequence in a separable Hilbert space contains a subsequence that is weakly convergent.

Proof- Suppose that $\|x_n\| \le M$ for all n and let $\{\phi_j\}$ denote a complete orthonormal family in H. Proceeding as we did in R^N, let $\{x_{n,k}\} \subset \{x_n\}$ denote a subsequence such that $\{(x_{n,k},\phi_j)_H\}$ is convergent (in R) for $j \le k$. Then for each j, $(x_{n,j},\phi_j)_H$ converges to a real limit a_j as n tends to infinity. It follows that the diagonal subsequence $\{x_{n,n}\}$ is such that $(x_{n,n},\phi_j)_H$ converges to a_j for $j \ge 1$. Now define

$$F(v) = Lim_n\ (x_{n,n},v)_H \quad \text{for} \quad \text{v in H.}$$

Then $\quad |F(v)| \le |\lim_n(x_{n,n},v)_H| \le M\|v\|_H$

from which it follows that F is a continuous linear functional on H. By the Riesz theorem, there exists an element, z_F in H such that $F(v) = (z_F,v)_H$ for all v in H.

But $\quad F(v) = F\left(\sum_i (v,\phi_i)_H\phi_i\right) = \lim_n(x_{n,n},\sum_i (v,\phi_i)_H\phi_i)_H$

$$= \sum_i \lim_n(x_{n,n},\phi_i)_H(v,\phi_i)_H = \sum_i a_i(v,\phi_i)_H \ ;$$

That is, $\quad F(v) = (z_F,v)_H = \sum_i a_i(v,\phi_i)_H$ for all v in H. Then by the Parseval-Plancherel identity, it follows that

$$z_F = \sum_i a_i\phi_i$$

and

$$(x_{n,n},v)_H \to (z_F,v)_H \quad \text{for all v in H.}\blacksquare$$

We should point out that a sequence $\{x_n\}$ is H is said to be **strongly bounded** if there is an M such that $\|x_n\|_H \le M$ for all n, and it is said to be **weakly bounded** if $|(x_n,v)_H| \le M$ for all n and all v in H. These two notions of boundedness coincide in a Hilbert space so it is sufficient to use the term bounded for either case.

Lemma 5- A sequence in a Hilbert space is weakly bounded if and only if it is strongly bounded.