LESSON 2 - 10 MARCH 2021

HILBERT SPACES

HILBERT $\subseteq$ BANACH $\subseteq$ COMPLETE METRIC $\subseteq$ TOPOLOGICAL

INNER PRODUCT SPACES Let H be a real vector space. A function $\langle \cdot , \cdot \rangle : H \times H \to \mathbb{R}$ is an INNER PRODUCT if

- $\langle x, y \rangle = \langle y, x \rangle \quad , \forall x, y \in H$ (Symmetric)
- $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle, \forall \lambda, \mu \in \mathbb{R}, x, y \in H$ (Bilinear)
- $\langle x, x \rangle \geq 0, \forall x$ and $\langle x, x \rangle = 0$ iff $x = 0$ (Positive definite)

The pair $(H, \langle \cdot , \cdot \rangle)$ is called an INNER PRODUCT SPACE

REMARK $(H, \langle \cdot , \cdot \rangle)$ inner product space. Then $\|x\| = \sqrt{\langle x, x \rangle}$ defines a norm on X.

CAUCHY-SCHWARZ INEQUALITY H inner prod space. Then $|\langle x, y \rangle| \leq \|x\| \|y\|$

HILBERT SPACE $(H, \langle \cdot , \cdot \rangle)$ inner product space. We say that H is a HILBERT SPACE if $(H, \| \cdot \|)$ is COMPLETE, with $\|x\| = \sqrt{\langle x, x \rangle}$

BASIS H Hilbert. A set of elements $\{e_n\}_{n \in \mathbb{N}} \subseteq H$ is called BASIS if

- $\langle e_i, e_j \rangle = 0$ for all $i \neq j$, $\langle e_i, e_i \rangle = 1 \quad \forall i$
- $\operatorname{span}\{e_j\}$ is dense in H

where for a set $A \subseteq H$ we define $\operatorname{span} A = \left\{ \sum_{j=1}^{n} \lambda_j x_j \mid \lambda_j \in \mathbb{R}, x_j \in A, n \in \mathbb{N} \right\}$

THEOREM If H is Hilbert separable then $\exists \{e_n\} \subseteq H$ basis

NOTATION Given a basis $\{e_n\}$ and $x \in H$ we define $x_k := \langle x, e_k \rangle$ k-th coordinate of x WRT $\{e_n\}$

PROPOSITION H Hilbert with basis $\{e_n\}$. Then

1) $\|v\|^2 = \sum_{j=1}^{+\infty} v_n^2$, $v_n := \langle v, e_n \rangle$

2) $\langle v, w \rangle = \sum_{j=1}^{+\infty} v_n w_n$

For a separable Hilbert space there is a natural correspondence between H and ℓ^2. Thus we can think of H as $\mathbb{R}^\infty$. To make this statement precise we need the following definition

(Recall: $\ell^2 = \{ (x_1, x_2, \ldots) \mid \sum_{j=1}^{+\infty} |x_j|^2 < +\infty \}$. This is normed by $\|x\| = \left(\sum_{j=1}^{+\infty} |x_j|^2 \right)^{1/2}$ and is Hilbert with inner product $\langle x, y \rangle = \sum_{j=1}^{+\infty} x_j y_j$)

DEF X normed space, $\{x_n\} \subseteq X$. We say that

$$\sum_{n=1}^{+\infty} x_n = x_0$$

if $S_n \to x_0$, where $S_n := \sum_{j=1}^{n} x_j$ partial sums.

THEOREM H Hilbert with basis $\{e_n\}$. Let $\{v_n\} \subseteq \mathbb{R}$. Then

$$\sum_{n=1}^{+\infty} v_n e_n \text{ converges in } H \iff \sum_{n=1}^{+\infty} v_n^2 \text{ converges in } \mathbb{R}$$

In particular $H \cong \ell^2$, with isomorphism $v \in H \mapsto (v_1, v_2, \ldots, v_n, \ldots) \in \ell^2$

Another nice aspect of Hilbert spaces is that $H = H^*$ dual space.

THEOREM (RIESZ) H Hilbert. Define the map $\Phi : H \to H^*$

$$\Phi(x)(z) := \langle x, z \rangle, \ \forall z \in H$$

Then Φ is invertible and $\|\Phi(x)\|_{H^*} = \|x\|_H$.

Thus $H \cong H^*$ isomorphic.

In particular H can be identified with H^*. Thus weak* and weak topologies coincide, and we can characterize weak convergence by

$$x_n \rightharpoonup x_0 \iff \langle x_n, z\rangle \to \langle x_0, z\rangle, \ \forall z \in H.$$

FURTHER PROPERTIES OF WEAK CONVERGENCE IN HILBERT

PROP H Hilbert with basis $\{e_n\}$. If $x_n \rightharpoonup x_0$ then

$$(x_n)_k \to (x_0)_k, \quad \forall k \in \mathbb{N}$$

WARNING We know that $x_n \rightharpoonup x_0$ does not imply $x_n \to x_0$. However it is also NOT true that $\|x_n\| \to \|x_0\|$ (i.e. the norm is not weakly continuous).

However, the following proposition relating strong convergence to weak convergence holds.

PROP H Hilbert. Then

$$x_n \to x_0 \iff x_n \rightharpoonup x_0 \text{ and } \|x_n\| \to \|x_0\|$$

NOTE: This is not saying that $\|x_n - x_0\| \to 0$

Another useful proposition is that weak convergence can be tested against a dense subset

PROP H Hilbert. Assume that $\{x_n\} \subseteq H$ is bounded, i.e. $\sup_n \|x_n\| < +\infty$. Suppose that $W \subseteq H$ is s.t $\overline{\operatorname{span} W} = H$ and

$$\langle x_n, w\rangle \to \langle x_0, w\rangle, \ \forall w \in W.$$

Then $x_n \rightharpoonup x_0$.

COROLLARY

H Hilbert with basis $\{e_n\}$. Let $\{x_n\}$ be BOUNDED. Then if

$$(x_n)_k \to (x_0)_k \quad , \forall k \in \mathbb{N}$$

we have $x_n \rightharpoonup x_0$.

[Proof: take $W = \{e_n\}$]

EXAMPLE

$$X = C[a,b] \quad \text{with} \quad \|u\|_\infty := \max_{x \in [a,b]} |u(x)|$$

$$Y = C^1[a,b] \quad \text{with} \quad \|u\|_1 := \|u\|_\infty + \|u'\|_\infty$$

Then $(X, \|\cdot\|_\infty)$ and $(Y, \|\cdot\|_1)$ are Banach spaces, but not Hilbert spaces.

Hint to show this: in an inner product space the parallelogram law holds:

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

$\forall x,y \in H$

1. CALCULUS IN NORMED SPACES

Reference : S. KESAVAN
"NONLINEAR FUNCTIONAL ANALYSIS,
A FIRST COURSE"
HINDUSTAN BOOK AGENCY, 2004

Throughout this section X, Y are real NORMED SPACES, $U \subseteq X$ is OPEN and $F: U \to Y$ is a given function.

GOAL: Construct a theory of differentiation for maps $F: U \to Y$

DEFINITION 1.1 We say that F is FRÉCHET DIFFERENTIABLE at $u_0 \in U$ if $\exists\, A_{u_0} \in \mathcal{L}(X,Y)$ s.t.

$$(*) \qquad \lim_{\|v\|_X \to 0} \frac{\| F(u_0+v) - F(u_0) - A_{u_0}(v) \|_Y}{\|v\|_X} = 0 .$$

REMARK If F is diff. at u_0 then $A_{u_0} \in \mathcal{L}(X,Y)$ satisfying $(*)$ is UNIQUE (Check it by exercise)

NOTATION If F is diff. at $u_0 \in U$ we call A_{u_0} the Fréchet derivative (or just derivative) of F at u_0. We denote

$$F'(u_0) := A_{u_0} \in \mathcal{L}(X,Y)$$

NOTE This generalizes diff. for maps $F: \mathbb{R}^n \to \mathbb{R}^m$. In this case the differential is $F'(u)(v) = DF(u)\,v$, with $DF(u) \in \mathbb{R}^{m \times n}$ matrix of partial derivatives of F, i.e.

$$[DF(u)]_{ij} = \frac{\partial F_i}{\partial x_j}(u), \qquad F = (F_1, \dots, F_m)$$

DEFINITION 1.2 Assume that F is diff. $\forall u_0 \in U$. Then we can define the map

$$F' : U \to \mathcal{L}(X,Y)$$
$$u \mapsto F'(u)$$

If F' is continuous we say that $F \in C^1(U,Y)$

WRT norm on X and operator norm on $\mathcal{L}(X,Y)$

In words we say F is continuously diff.

EXAMPLES The most common examples throughout the course will be real valued functions, i.e., $Y = \mathbb{R}$

1) X normed, $U \subseteq X$ open, $F : U \to \mathbb{R}$. If F is diff at $u_0 \in U$ then $F'(u_0) \in \mathcal{L}(X,\mathbb{R}) = X^*$
Then if F diff in U, the derivative defines an application $F' : U \to X^*$ $(u \mapsto F'(u) \in X^*)$

2) H Hilbert, $U \subseteq H$ open, $F : U \to \mathbb{R}$. If F is diff. at $u_0 \in U$ then $F'(u_0) \in \mathcal{L}(H,\mathbb{R})$. By Riesz's Thm $\exists!$ $z_0 \in H$ s.t.

$$F'(u_0)(w) = \langle z_0, w \rangle \quad , \forall w \in H$$

We denote $z_0 := \text{grad}\, F(u_0)$ (gradient of F at u_0).

PROPOSITION 1.3 Assume that F is diff at $u_0 \in U$. Then F is continuous at u_0.

Proof Introduce the notation $o(\|v\|_X)$ for a quantity such that

$\frac{o(\|v\|_X)}{\|v\|_X} \to 0$ as $\|v\|_X \to 0$. Since U is open and $u_0 \in U$ then $\exists \varepsilon > 0$ such that $B_\varepsilon(u_0) \subseteq U$. Let $v \in B_\varepsilon(0)$ so that $u_0 + v \in B_\varepsilon(u_0) \subseteq U$: then

$$\|F(u_0+v) - F(u_0)\|_Y \leq \|F(u_0+v) - F(u_0) - A_{u_0}(v)\|_Y + \|A_{u_0}(v)\|_Y$$

$$\leq o(\|v\|_X) + \|A_{u_0}\|_{\mathcal{L}(X,Y)} \|v\|_X \quad (\text{since } A_{u_0} \in \mathcal{L}(X,Y))$$

$$= o(\|v\|_X) \to 0 \text{ as } \|v\|_X \to 0 \qquad \square$$

THEOREM 1.4 (CHAIN RULE) X, Y, Z normed, $U \subseteq X$, $V \subseteq Y$ open, $F: U \to V$, $G: V \to Z$. Assume F is diff at $u_0 \in U$ and G is diff. at $v_0 := F(u_0) \in V$. Then $G \circ F : U \to Z$ is diff. at u_0 with

$$(G \circ F)'(u_0) = G'(v_0) \circ F'(u_0) \in \mathcal{L}(X, Z)$$

composition of linear continuous operators in $\mathcal{L}(X,Y)$ and $\mathcal{L}(Y,Z)$

The proof is very simple, and we thus omit it. If you are interested you can find it in the book of KESAVAN, PROPOSITION 1.1.1 page 7

DEFINITION 1.5 We say that F is GÂTEAUX DIFFERENTIABLE at $u_0 \in U$ in the DIRECTION $v \in X$ if

$$F'_g(u_0)(v) := \lim_{t \to 0} \frac{F(u_0 + tv) - F(u_0)}{t} \in Y,$$

Here $t \in \mathbb{R}$

i.e., if the above limit exists.

NOTE The Gâteaux derivative (G. derivative) generalizes the directional derivative for maps $F: \mathbb{R}^n \to \mathbb{R}$.
In this case we have

$$F'_G(u)(v) = \frac{\partial F}{\partial v}(u) = \nabla F(u_0) \cdot v$$

If F diff at u

Scalar product in $\mathbb{R}^n$

WARNING $F'_G(u): X \to Y$ is always linear. But in general we don't have that $F'_G(u) \in \mathcal{L}(X,Y)$, as it happens for Fréchet derivatives.

PROPOSITION 1.6 If F diff. at $u_0 \in U$ then F is G. diff at u_0 in every direction v and

$$F'_G(u_0)(v) = F'(u_0)(v)$$

Gâteaux derivative can be computed by Frechet derivative, and viceversa

Proof

$$\frac{F(u_0+tv) - F(u_0)}{t} =$$

$$= \underbrace{\frac{F(u_0+tv) - F(u_0) - F'(u_0)(tv)}{t\|v\|_X}\|v\|_X}_{\to 0 \text{ as } t \to 0,} + \frac{F'(u_0)(tv)}{t}$$

since F is diff. at u_0

(note this is converging in Y to 0)

$$= o(t) + F'(u_0)(v) \to F'(u_0)(v) \text{ as } t \to 0$$

□

Here we used that $F'(u_0)$ is lin. operator, so $F'(u_0)(tv) = tF'(u_0)(v)$

WARNING The converse of Prop 1.6 does not hold, i.e.

$$\text{Gâteaux Diff.} \not\Rightarrow \text{Fréchet Diff.}$$

For example take $F : \mathbb{R}^2 \to \mathbb{R}$

$$F(x,y) := \begin{cases} \dfrac{x^5}{(y-x^2)^2+x^4} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

It is easy to check that $F'_G(0)(v) = 0$, $\forall v \in \mathbb{R}$. So F is G-diff at $0 = (0,0)$ in every direction. Thus, if F was Fréchet diff we would have (by Prop 1.6) that $F'(0)(v) = 0$. Thus by def of derivative

$$(*) \qquad \lim_{\|v\|_{\mathbb{R}^2} \to 0} \frac{|F(v) - F(0)|}{\|v\|_{\mathbb{R}^2}} = 0$$

However if in the above limit we consider $v \in \{y = x^2\}$ we obtain

$$\lim_{\substack{\|v\|_{\mathbb{R}^2} \to 0 \\ v \in \{y = x^2\}}} \frac{|F(v) - F(0)|}{\|v\|_{\mathbb{R}^2}} = 1,$$

which contradicts $(*)$. Thus F is not Fréchet diff at 0.

REMARK Proposition 1.6 is very useful to guess the Fréchet derivative of a function $F : U \to Y$, as the Gâteaux derivative can be computed via a formula

EXAMPLE $X = C[0,1]$, with norm $\|\cdot\|_\infty$. Define $F: X \to \mathbb{R}$ by

$$F(u) = \int_0^1 \sin(u(x))\, dx$$

What could be the derivative of F? Let us compute the gâteaux derivative at u in the direction v:

$$\lim_{t\to 0} \frac{F(u+tv) - F(u)}{t} = \lim_{t\to 0} \int_0^1 \frac{\sin(u+tv) - \sin(u)}{t}\, dx$$

From the chain Rule (THEOREM 1.4) we have:

$$\lim_{t\to 0} \frac{\sin(u(x) + tv(x)) - \sin(u(x))}{t} = \cos(u(x))\, v(x) \quad , \text{ uniformly in } x \in [0,1]$$

Thus we can pass the limit under the integral and obtain

$$F_g'(u)(v) = \int_0^1 \cos(u(x))\, v(x)\, dx . \qquad \left(F_g'(u) \text{ is linear!}\right)$$

If F is Fréchet diff. then by PROPOSITION 1.6 we must have $F'(u) = F_g'(u)$. So we guess that $F_g'(u)$ is the Fréchet derivative of F at u. Indeed notice that $F_g'(u) \in \mathcal{L}(X, \mathbb{R})$, since

$$|F_g'(u)(v)| \leq \|v\|_\infty \int_0^1 |\cos(u(x))|\, dx \leq \|v\|_\infty \Rightarrow \sup_{\|v\|_X \leq 1} |F_g(u)(v)| < +\infty .$$

Moreover it is easy to see that

$$\lim_{\|v\|_X \to 0} \frac{|F(u+v) - F(u) - F_g'(u)(v)|}{\|v\|_X} = 0$$

showing that F is Fréchet diff. at u with $F'(u)(v) = \int_0^1 \cos(u(x))\, v(x)\, dx$.

QUESTION Assume that $F: U \to Y$ is gâteaux differentiable. Under which assumptions on F are we guaranteed Fréchet differentiability?

To answer the above question, we need the following theorem.

THEOREM 1.7 (MEAN VALUE)

Suppose F is G. diff. in U in every direction. Let $x_1, x_2 \in U$ be such that the segment

$$[x_1, x_2] := \{ x_1 + t(x_2 - x_1),\ t \in [0,1] \} \subseteq U$$

Assume also that $F_g'(u) : X \to Y$ is s.t. $F_g'(u) \in \mathcal{L}(X,Y)$, $\forall u \in U$. Then

$$\| F(x_2) - F(x_1) \|_Y \leq \sup_{u \in [x_1, x_2]} \| F_g'(u) \|_{\mathcal{L}(X,Y)} \| x_2 - x_1 \|_X$$

Proof: If $F(x_1) = F(x_2)$ then the thesis is trivial. Thus assume that $F(x_1) \neq F(x_2)$. We now employ the following:

FACT (COROLLARY of HAHN-BANACH) Y normed space, $z \in Y$, $z \neq 0$. Then $\exists\, \Lambda \in Y^*$ s.t.

$$\| \Lambda \|_{Y^*} = 1 \quad \text{and} \quad \Lambda(z) = \| z \|_Y .$$

Thus let $\Lambda \in Y^*$ be such that $\Lambda(F(x_2) - F(x_1)) = \| F(x_2) - F(x_1) \|_Y$ and $\| \Lambda \|_{Y^*} = 1$.

Define the segment function $\alpha : [0,1] \to U$ by $\alpha(t) := x_1 + t(x_2 - x_1)$.

Consider the map $H : [0,1] \to \mathbb{R}$ defined by

$$H := \Lambda \circ F \circ \alpha .$$

The thesis is obtained by applying the classical Mean Value Theorem to H. Thus, all we need to show is that H is differentiable.

WARNING: It would be tempting to apply the Chain Rule of Theorem 1.4 to H. However F is only Gâteaux differentiable, and in general the Chain Rule does not apply in this case.

We will check by hand that H is differentiable. Thus let $t \in [0,1]$, and $\tau \neq 0$ be such that $(t+\tau) \in [0,1]$. We have

$$(*) \qquad \frac{H(t+\tau) - H(t)}{\tau} = \Lambda \left[\frac{F(\alpha(t+\tau)) - F(\alpha(t))}{\tau} \right] \qquad \text{(being } \Lambda \text{ linear)}$$

Note that $F(\alpha(t+\tau)) = F(\alpha(t) + \tau (x_2 - x_1))$. Since F is gâteaux diff. at $\alpha(t)$ we then get

$$\frac{F(\alpha(t+\tau)) - F(\alpha(t))}{\tau} = \frac{F(\alpha(t) + \tau(x_2 - x_1)) - F(\alpha(t))}{\tau} \longrightarrow F'_g(\alpha(t))(x_2 - x_1)$$

as $\tau \to 0$. Note that by definition the above convergence is WRT the norm of Y. As $\Lambda \in Y^*$ is continuous, by taking the limit as $\tau \to 0$ in $(*)$ we get

$$H'(t) = \lim_{\tau \to 0} \frac{H(t+\tau) - H(t)}{\tau} = \Lambda \left[F'_g(\alpha(t))(x_2 - x_1) \right].$$

In particular H is diff. in $[0,1]$. Therefore we can apply the Mean Value Theorem to find $\xi \in (0,1)$ such that

$$(**) \qquad H(1) - H(0) = H'(\xi) \qquad \left(= H'(\xi)(1-0) \right)$$

Now

$$\begin{aligned} H(1) - H(0) &= \Lambda[F(\alpha(1))] - \Lambda[F(\alpha(0))] \\ &= \Lambda[F(x_2)] - \Lambda[F(x_1)] \\ &= \Lambda[F(x_2) - F(x_1)] \\ &= \| F(x_2) - F(x_1) \|_Y \end{aligned}$$

by the properties of Λ.

On the other hand, as we computed, we have

$$H'(\xi) = \Lambda\left[F_g'(\alpha(\xi))\,(x_2 - x_1) \right]$$

and so

$$|H'(\xi)| \leq \|\Lambda\|_{Y^*} \, \|F_g'(\alpha(\xi))\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X$$

(Using that Λ and $F_g'(\alpha(\xi))$ are linear and bounded)

$$\leq \sup_{u \in [x_1, x_2]} \|F_g'(u)\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X$$

(Using that $\|\Lambda\|_{Y^*} = 1$ and that $\alpha(\xi) \in [x_1, x_2]$)

From (**) we then get

$$\|F(x_2) - F(x_1)\|_Y = H(1) - H(0) = H'(\xi)$$

$$\leq \sup_{u \in [x_1, x_2]} \|F_g'(u)\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X \qquad \square$$

COROLLARY 1.8 (of MEAN VALUE) Make the same assumptions of Theorem 1.7. Then

$$\|F(x_2) - F(x_1) - F_g'(x_1)(x_2 - x_1)\|_Y \leq \sup_{u \in [x_1, x_2]} \|F_g'(u) - F_g'(x_1)\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X$$

Proof Define $H : U \to Y$ by $H(u) := F(u) - F_g'(x_1)(u)$. Note that

$F_g'(x_1) \in \mathcal{L}(X, Y)$ by assumption. In particular $F_g'(x_1) : X \to Y$ is Fréchet differentiable with derivative constantly equal to itself.

(Check it by exercise: X, Y normed spaces, $T \in \mathcal{L}(X,Y)$. Then T is Fréchet diff with $T'(u) = T$, $\forall u \in X$).

Therefore H is Gâteaux diff. with $H_g'(u) = F_g'(u) - F_g'(x_1)$. Thus $H_g'(u) \in \mathcal{L}(X,Y)$ for all $u \in U$. Thus H satisfies assumptions of THEOREM 1.7. Applying THM 1.7 to H we conclude. $\square$

We are finally ready to answer our question:

"When does Gâteaux diff. imply Fréchet diff.?"

THEOREM 1.9 Assume that $F: U \to Y$ is Gâteaux diff. at every point of U and in every direction. Also suppose that $F'_g(u) \in \mathcal{L}(X,Y)$ for all $u \in U$. Define the map

$$F'_g : U \longrightarrow \mathcal{L}(X,Y)$$
$$u \longmapsto F'_g(u)$$

If F'_g is continuous at u_0 then F is Fréchet diff at u_0 and

continuity is WRT norm on X and operator norm on $\mathcal{L}(X,Y)$

$$F'(u_0)(v) = F'_g(u_0)(v), \quad \forall v \in X$$

Proof Apply COROLLARY 1.8 to points $x_1 := u_0$, $x_2 := u_0 + v$. Since U is open, for v s.t. $\|v\|_X$ is sufficiently small we have $[x_1, x_2] \subseteq U$. By COROLLARY 1.8 we have

$$(*) \quad \| F(u_0 + v) - F(u_0) - F'_g(u_0)(v) \|_Y \leq \sup_{u \in [x_1, x_2]} \| F'_g(u) - F'_g(u_0) \|_{\mathcal{L}(X,Y)} \|v\|_X$$

Recall that $[x_1, x_2] = [u_0, u_0 + v]$. As F'_g is continuous at u_0 we have

$$\sup_{u \in [x_1, x_2]} \| F'_g(u) - F'_g(u_0) \|_{\mathcal{L}(X,Y)} \longrightarrow 0 \quad \text{as} \quad \|v\|_X \to 0$$

(in practice this implies $[x_1, x_2] \to \{u_0\}$)

Therefore, dividing $(*)$ by $\|v\|_X$ and taking the limit as $\|v\|_X \to 0$ concludes.

$\square$