Lecture 13: Fréchet Derivative. Hahn-Banach Theorems and Applications.

Recall: An operator $F : X \to Y$ (between normed linear spaces X and Y) is Fréchet differentiable at $a \in X$ if and only if there exists a bounded linear operator $DF(a) : X \to Y$ such that

$$\lim_{h\to 0} \frac{\|F(a+h) - F(a) - DF(a)h\|}{\|h\|} = 0.$$

If F is Fréchet differentiable, we can write

$$F(a+h) = F(a) + DF(a)h + R(a,h), \quad \text{where} \quad \lim_{h\to 0} \frac{\|R(a,h)\|}{\|h\|} = 0,$$

or

$$F(a+h) = F(a) + DF(a)h + \|h\| r(h), \quad \text{where} \quad \lim_{h\to 0} r(h) = 0,$$

or

$$F(a+h) = F(a) + DF(a)h + o(\|h\|).$$

Remark 2. *1. The Fréchet derivative is a generalization of derivative in $\mathbb{R}$. That is, if $F : \mathbb{R} \to \mathbb{R}$ is Fréchet differentiable at $a \in \mathbb{R}$, then F is differentiable at a and*

$$DF(a)(x) = F'(a)x, \quad \forall\, x \in \mathbb{R}.$$

Proof. Since F is Fréchet differentiable at $a \in \mathbb{R}$, there exists $DF(a) \in B(\mathbb{R}, \mathbb{R})$ such that

$$0 = \lim_{h\to 0} \frac{|F(a+h) - F(a) - DF(a)(h)|}{|h|}$$

Since

$$B(\mathbb{R}, \mathbb{R}) = \{L : \mathbb{R} \to \mathbb{R} \ s.t.\ L(x) = cx,\ \forall x \in \mathbb{R},\ \text{for some } \ c \in \mathbb{R}\},$$

there exists some constant $c \in \mathbb{R}$ such that

$$DF(a)(x) = cx, \forall\, x \in \mathbb{R}.$$

So

$$0 = \lim_{h\to 0} \frac{|F(a+h) - F(a) - DF(a)(h)|}{|h|} = \lim_{h\to 0} \frac{|F(a+h) - F(a) - ch|}{|h|} = \lim_{h\to 0} \left| \frac{F(a+h) - F(a)}{h} - c \right|.$$

So

$$c = \lim_{h\to 0} \frac{F(a+h) - F(a)}{h},$$

which implies F is differentiable at a and $c = F'(a)$. Therefore,

$$DF(a) : \mathbb{R} \to \mathbb{R}, \quad DF(a)(x) = F'(a)x, \quad \forall\, x \in \mathbb{R}.$$

□

2. *To compute the Fréchet-derivative of an operator $F : X \to Y$ at $a \in X$, where X and Y are normed linear spaces, we write $F(a+h)-F(a)$ as a summation of a linear operator (w.r.t. h) and a remainder (which is nonlinear in h)*
$$F(a+h) - F(a) = Lh + R(a,h),$$
and prove that L is bounded and
$$\lim_{h\to 0} \frac{\|R(a,h)\|}{\|h\|} = 0.$$
The linear operator $L \in B(X,Y)$ is the $DF(a)$ in the definition.

3. *If $F : X \to Y$ is Fréchet-derivative at $a \in X$, where X and Y are normed linear spaces, then for any $x \in X$, we have*
$$DF(a)(x) = \lim_{t\to 0} \frac{F(a+tx) - F(a)}{t}, \quad t \in \mathbb{R}. \quad \text{(Prove this)}$$
Note: This formula is used to compute the Fréchet-derivative of an operator F. After this, we need to check that $DF(a)$ is a bounded linear operator and $R(a,h) = F(a+h) - F(a) - DF(a)h$ satisfies
$$\lim_{h\to 0} \frac{\|R(a,h)\|}{\|h\|} = 0.$$
Here is an example of an operator $F : X \to Y$, where $\lim_{t\to 0} \frac{F(a+tx) - F(a)}{t}$, $t \in \mathbb{R}$ exists for an $a \in X$ but F is not Fréchet differentiable at a. Consider $F : \mathbb{R}^2 \to \mathbb{R}$ given by
$$F(x_1, x_2) = \begin{cases} 0 & \text{if} \quad x_2 = 0, \\ \dfrac{x_1^3}{x_2} & \text{if} \quad x_2 \neq 0. \end{cases}$$
The operator F is not continuous at $(0,0)$, for example, $F(t,t^3) \to 1$ as $t \to 0$, but $F(0,0) = 0$. On the other hand,
$$\frac{F(tx) - F(0,0)}{t} = \begin{cases} 0 & \text{if} \quad x_2 = 0, \\ t\dfrac{x_1^3}{x_2} & \text{if} \quad x_2 \neq 0. \end{cases}$$
So $\lim_{t\to 0} \frac{F(tx) - F(0,0)}{t} = 0$ for any $(x_1, x_2) \in \mathbb{R}^2$.

Example 2. 1. *Let $X = C_0^1[0,1]$ be the space of all C^1 functions on $[0,1]$ which vanish at the endpoints with norm*
$$\|u\| = \left[\int_0^1 \left[u^2 + (u')^2 \right] dx \right]^{1/2}.$$
Consider an operator $K : X \to \mathbb{R}$ defined by
$$K(u) = \int_0^1 \left[u^3 + (u')^2 \right] dx.$$
Compute the Fréchet derivative of K.

2. *Let $X = C[a,b]$ with $\|\cdot\|_\infty$ norm. Let $T : X \to X$ be the nonlinear integral operator defined by*

$$(Tu)(x) = u(x)\int_a^b K(x,s)u(s)\,ds,$$

where $K(x,s)$ is continuous on $[a,b]\times[a,b]$. Compute the Fréchet derivative of K.

Proof. Exercise and See Dr. Vrscay's notes (attached in the next pages). □

Proposition 19. *Let X and Y be normed linear spaces over $\mathbb{K}$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). If F is Fréchet differentiable at $a \in X$ then F is continuous at a.*

Proof. Since F is Fréchet differentiable at $a \in X$, there exists $\delta > 0$ such that when $\|h\| < \delta$, we have

$$\|F(a+h) - F(a) - DF(a)(h)\| \leq \|h\|.$$

So for all $h \in X$ with $\|h\| < \delta$, we have

$$\|F(a+h)-F(a)\| \leq \|F(a+h)-F(a)-DF(a)(h)\|+\|DF(a)(h)\| \leq \|h\|+\|DF(a)\|\|h\| = (1+\|DF(a)\|)\|h\|$$

As $h \to 0$, $(1+\|DF(a)\|)\|h\| \to 0$. Therefore,

$$\lim_{h\to 0}\|F(a+h) - F(a)\| = 0,$$

which means $\lim_{h\to 0} F(a+h) = F(a)$. Therefore F is continuous at a. □

Proposition 20. *Let X, Y and Z be normed linear spaces over $\mathbb{K}$.*

1. *Let $f, g : X \to Y$ be Fréchet derivative at $a \in X$. Then for any $\alpha, \beta \in \mathbb{K}$, we have*

$$D(\alpha f + \beta g)(a) = \alpha Df(a) + \beta Dg(a).$$

2. *(Chain Rule) Suppose $F : X \to Y$ is Fréchet differentiable at $a \in X$, $G : Y \to Z$ is Fréchet differentiable at $F(a)$. Then $G \circ F : X \to Z$ is Fréchet differentiable at a and*

$$D(G\circ F)(a) = DG(F(a))\,DF(a).$$

Proof. Set $b = F(a)$. By the assumptions,

$$F(a+h) - F(a) = DF(a)h + \|h\|r_1(h)$$
$$G(b+k) - G(b) = DG(b)k + \|k\|r_2(k),$$

where $\|r_1(h)\| \to 0$ as $h \to 0$ and $\|r_2(k)\| \to 0$ as $k \to 0$, for $h \in X, k \in Y$.
For $h \in X$, denote $k = DF(a)h + \|h\|r_1(h)$. Now we compute

$$\begin{aligned} G(F(a+h)) - G(F(a)) &= G\Big(b + DF(a)h + \|h\|r_1(h)\Big) - G(b) \\ &= DG(b)\Big(DF(a)h + \|h\|r_1(h)\Big) + \|k\|r_2(k) \\ &= DG(b)DF(a)h + \|h\|DG(b)\Big(r_1(h)\Big) + \|k\|r_2(k) \end{aligned}$$

The operator $DG(F(a))DF(a) : X \to Z$ is a bounded linear operator since it is the composition of two bounded linear operators $DF(a)$ and $DG(F(a))$. Observe that

$$\|k\| \leq \Big(\|DF(a)\| + \|r_1(h)\|\Big)\|h\|, \quad \forall h \in X, \tag{12}$$

and

$$\frac{\|r(h)\|}{\|h\|} \leq \|DG(b)\|\|r_1(h)\| + \frac{\|k\|}{\|h\|}\|r_2(k)\| \leq \|DG(b)\|\|r_1(h)\| + \Big(\|DF(a)\| + \|r_1(h)\|\Big)\|r_2(k)\| \tag{13}$$

Now letting $h \to 0$. Since $\|r_1(h)\| \to 0$ as $h \to 0$, from (12), we have $k \to 0$ and hence $r_2(k) \to 0$. Therefore, from (13), we have

$$\frac{\|r(h)\|}{\|h\|} \to 0 \quad \text{as} \quad h \to 0.$$

□

2.14 Hahn-Banach Theorems. Generalized Mean Value Theorem. Separation Theorems.

Definition 1. *Let X be a vector space over $\mathbb{K}$. We say that $p : X \to [0, \infty)$ is sublinear if it satisfies*

$$p(\lambda x) = \lambda p(x) \quad \textit{(positive homogeneous)},$$
$$p(x + y) \leq p(x) + p(y) \quad \textit{(triangle inequality)},$$

for any $x, y \in X$ and real $\lambda > 0$.

Lemma 6 (Zorn's Lemma)**.** *Suppose S is a nonempty, partially ordered set (reflexivity, antisymmetry, and transitivity). If every totally ordered subset C (that is, every two elements in C are comparable) of S has an upper bound; that is, there is some $u \in S$ such that*

$$x \leq u \quad \textit{for all} \quad x \in C.$$

Then S has at least one maximal element; that is there is some $m \in S$ such that for any $x \in S$,

$$\textit{if} \quad m \leq x \quad \textit{then} \quad m = x.$$

Theorem 1 (The Hahn-Banach Theorem for linear spaces)**.** *Let X_0 be a subspace of a real vector space X and p be a sublinear on X. If $f_0 : X_0 \to \mathbb{R}$ is a linear functional such that*

$$f_0(x) \leq p(x), \quad \forall\, x \in X_0,$$

then there is a linear functional $f : X \to \mathbb{R}$ such that

$$f|_{X_0} = f_0, \quad \textit{(i.e., f is a linear extension of f_0)}$$

and

$$f(x) \leq p(x), \quad \forall\, x \in X.$$

Sketch of the proof. **Step 1:** We first prove the statement in the special case when $X = X_0 + \text{span}(v)$ with a fixed $v \notin X_0$. Set

$$f(\lambda v + x_0) = f_0(x_0) + \lambda c, \quad \forall\, x_0 \in X_0,\ \forall \lambda \in \mathbb{R},$$

where $c \in \mathbb{R}$ is a fixed number satisfying

$$\sup_{u \in X_0} (f_0(u) - p(u - v)) \leq c \leq \inf_{w \in X_0} (p(w + v) - f_0(w)).$$

We first show that such c exists. Indeed, for all $u, w \in X_0$, we have

$$f_0(u) + f_0(w) = f_0(u + w) \leq p(u + w) \leq p(u - v) + p(w + v).$$

Therefore,

$$f_0(u) - p(u - v) \leq p(w + v) - f_0(w), \quad \forall u, w \in X_0,$$

which means such c exists. Next, we will verify that the defined f is a linear functional on X , $f|_{X_0} = f_0$ (leave it as an exercise).

Finally, we will prove that $f(\lambda v + x_0) \leq p(\lambda v + x_0)$ for all $x_0 \in X_0$ and for all $\lambda \in \mathbb{R}$. The statement is true for $\lambda = 0$. For $\lambda > 0$, from the requirement on c,

$$c \leq p(w + v) - f_0(w), \quad \forall w \in X_0$$

we have

$$c \leq p(\lambda^{-1} x_0 + v) - f_0(\lambda^{-1} x_0) = \lambda^{-1}\Big(p(x_0 + \lambda v) - f_0(x_0)\Big),$$
$$f(x_0 + \lambda v) = \lambda c + f_0(x_0) \leq p(x_0 + \lambda v).$$

Similarly, for $\lambda < 0$, from the requirement on c,

$$c \geq f_0(u) - p(u - v), \quad \forall u \in X_0,$$

we have

$$c \geq f_0(-\lambda^{-1} x_0) - p(-\lambda^{-1} x_0 - v) = -\lambda^{-1}(f_0(x_0) - p(x_0 + \lambda v)),$$
$$f(x_0 + \lambda v) = \lambda c + f_0(x_0) \leq p(x_0 + \lambda v).$$

Step 2: Let $\mathcal{S}$ be the set of all linear extensions g of f_0 defined on a vector space $X_g \subset X$ and satisfying the property $g(x) \leq p(x)$ for all $x \in X_g$. Since $f_0 \in \mathcal{S}$, $\mathcal{S}$ is not empty. Define a partial ordering on S by $g \leq h$ means that h is a linear extension of g. For any totally ordered subset $\mathcal{C} \subset \mathcal{S}$, let

$$Y = \bigcup_{g \in \mathcal{C}} X_g, \quad g_{\mathcal{C}}(x) = g(x) \quad \text{for any } g \in \mathcal{C} \text{ such that } x \in X_g.$$

Since $\mathcal{C}$ is totally ordered, $g_{\mathcal{C}}$ is well-defined. Moreover $g_{\mathcal{C}} \in \mathcal{S}$ and is an upper bound for $\mathcal{C}$. Applying the Zorn's lemma, $\mathcal{S}$ has at least one maximal element f. By definition, f is a linear extension of f_0 and $f(x) \leq p(x)$ for all $x \in X_f$. It remains to show that $X_f = X$. If not, there exists $v \in X \setminus X_f$. Applying results from Step 1, we can construct a linear extension of f to $\widetilde{f}$ on $X_f + \mathbb{R}v$. This contradicts the maximality of f. Therefore, $X_f = X$, which completes the proof. □

Theorem 2 (The Hahn-Banach Theorem for normed spaces)**.** *Let X_0 be a subspace of a normed space X over $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Let $f_0 : X_0 \to \mathbb{R}$ be a linear functional such that*

$$|f_0(x)| \leq \alpha \|x\| \quad \forall\, x \in X_0 \text{ and fixed } \ \alpha \geq 0.$$

Then there is a linear functional $f : X \to \mathbb{R}$ such that

$$f|_{X_0} = f_0, \quad \text{(i.e., } f \text{ is a linear extension of } f_0\text{)}$$

and

$$|f(x)| \leq \alpha \|x\| \quad \forall\, x \in X.$$

Sketch of the proof. We prove the case $\mathbb{K} = \mathbb{R}$. Define

$$p(x) := \alpha \|x\| \quad \forall x \in X.$$

We can verify that $p(x)$ is sublinear (Prove this). Since $f_0(x) \leq |f_0(x)| \leq p(x)$, by the Hahn-Banach theorem for linear spaces, there is a linear functional $f : X \to \mathbb{R}$ such that $f|_{X_0} = f_0$ and $f(x) \leq \alpha \|x\|$. Since

$$-f(x) = f(-x) \leq \alpha \| - x\| = \alpha \|x\|,$$

we also have $f(x) \geq -\alpha \|x\|$ for all $x \in X$. Therefore, $|f(x)| \leq \alpha \|x\|$ for all $x \in X$. □