ANALYSIS 2
RECITATION SESSION OF WEEK 3

JACOB SHAPIRO

1. Differentiation in Banach Spaces–The Fréchet and Gâteaux derivatives

Following [1] (which deals only with the Banach spaces $\mathbb{R}^n$, whereas we generalize the definitions), we define the concept of differentiability in Banach spaces. In what follows, X and Y denote two Banach spaces and $E \in \mathrm{Open}(X)$. Furthermore, $f \in Y^E$ and $x \in E$.

1.1. **Definition.** (Fréchet derivative) f is called differentiable at x iff $\exists$ a linear map $A \in Y^X$ such that

$$\lim_{h\to 0_X} \frac{\|f(x+h) - f(x) - A(h)\|_Y}{\|h\|_X} = 0 \tag{1}$$

in which case we write $f'(x) = A$.

1.2. **Example.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto (x_1)^2 + (x_2)^2$. Then f is differentiable for every $x \in \mathbb{R}^2$ and $f'(x) : \mathbb{R}^2 \to \mathbb{R}$ is given by $\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \mapsto \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$ or simply $f'(x) = \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}$. To see this, calculate the limit:

$$\begin{aligned}
\lim_{h\to 0_{\mathbb{R}^2}} \frac{\|f(x+h) - f(x) - A(h)\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} &= \lim_{h\to 0_{\mathbb{R}^2}} \frac{\left\|(x_1+h_1)^2 + (x_2+h_2)^2 - (x_1)^2 - (x_2)^2 - \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}\right\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} \\
&= \lim_{h\to 0_{\mathbb{R}^2}} \frac{\|0\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} \\
&= 0
\end{aligned}$$

1.3. *Remark.* Observe that if equation (1) holds for both A and B then $A = B$ ([1] Theorem 9.12).

1.4. *Remark.* Observe that f' defines a map from $E \to \mathcal{L}(X, Y)$, $x \mapsto (h \mapsto A(h) \forall h \in X)$. As E and $\mathcal{L}(X, Y)$ are both Banach spaces, we may ask what is the derivative of this map. It turns out that the derivative is just A again:

$$\lim_{h\to 0_X} \frac{\|A(x+h) - A(x) - A(h)\|_Y}{\|h\|_X} = 0$$

by linearity of A.

1.5. *Claim.* (Remark 9.13 (c) in [1]) If f is differentiable at x then f is continuous at x.

1.6. **Definition.** (Gâteaux derivative) If for some $h \in X$ and $t \in \mathbb{R}$ the limit

$$\lim_{t\to 0} \frac{f(x+th) - f(x)}{t}$$

exists then we define $(\partial_h f)(x) := \lim_{t\to 0} \frac{f(x+th)-f(x)}{t}$ and call it the partial (or Gâteaux) derivative of f at x in the direction defined by h. Note that $[(\partial_h f)(x)] \in Y$. We also say that f is Gâteaux differentiable in the direction of h at x.

1.7. *Claim.* If $f : X \to Y$ is differentiable at $x \in X$ then it is Gâteaux differentiable in any direction $h \in X$ and we have $(f'(x))(h) = (\partial_h f)(x)$. (Theorem 9.17 in [1]).

To be concrete, if $f : \mathbb{R}^n \to \mathbb{R}^m$ then

$$f'(x) = h \mapsto \begin{bmatrix} (\partial_{\hat{e}_1}(f\cdot\hat{e}_1))(x) & \dots & \dots & (\partial_{\hat{e}_n}(f\cdot\hat{e}_1))(x) \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (\partial_{\hat{e}_1}(f\cdot\hat{e}_m))(x) & \dots & \dots & (\partial_{\hat{e}_n}(f\cdot\hat{e}_m))(x) \end{bmatrix} h$$

and in our example above 1.2, we can compute

$$\begin{aligned}
f'(x) &= \begin{bmatrix} (\partial_{\hat{e}_1} f)(x) & (\partial_{\hat{e}_2} f)(x) \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}
\end{aligned}$$

so that we see we can think of $\partial_{\hat{e}_j} f$ can be thought of as the ordinary derivative of f (from Analysis 1) as if it only depended on x_j and all other variables of it are constant. This also gives you a "recipe" to compute $f'(x)$ using the partial derivatives, which you know how to compute, from methods of Analysis 1.

Date: 2nd of March 2015.

1.8. **Example.** (Problem 9.6 in [1]) Define $f : \mathbb{R}^2 \to \mathbb{R}$ by $\begin{cases} 0 & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \frac{x_1 x_2}{(x_1)^2 + (x_2)^2} & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{cases}$.

1.9. *Claim.* f is not continuous at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and so, by 1.5 f is not differentiable.

Proof. For f to be continuous at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ we need to have that $\forall \varepsilon > 0 \; \exists$ some $\delta > 0$ such that if $\|x\|_{\mathbb{R}^2} < \delta$ then $\|f(x) - f(0)\|_{\mathbb{R}^1} < \varepsilon$. Explicitly, if $\sqrt{(x_1)^2 + (x_2)^2} < \delta$ then $\left|\frac{x_1 x_2}{(x_1)^2 + (x_2)^2}\right| < \varepsilon$. This can be easily seen to be impossible because

$$\begin{aligned} \left|\frac{x_1 x_2}{(x_1)^2 + (x_2)^2}\right| &= \left|\frac{x_1}{(x_1)^2 + (x_2)^2} \frac{x_2}{(x_1)^2 + (x_2)^2}\right| \\ &= |\hat{x}_1| |\hat{x}_2| \end{aligned}$$

if we pick $\varepsilon = \frac{1}{4}$ because $|\hat{x}_1||\hat{x}_2| = |\cos(\theta)\sin(\theta)|$ where θ is the angle between x and the $\hat{e}_1$, and, in general, $|\cos(\theta)\sin(\theta)| \in [0, \frac{1}{2}]$. □

1.10. *Claim.* $\partial_1 f$ and $\partial_2 f$ exist for every point $x \in \mathbb{R}^2$.

Proof. At any point $x \neq 0$ we have

$$\begin{aligned} (\partial_1 f)(x) &= \partial_1 \frac{x_1 x_2}{(x_1)^2 + (x_2)^2} \\ &= \frac{x_2 \left[(x_1)^2 + (x_2)^2\right] - [x_1 x_2][2x_1]}{\left[(x_1)^2 + (x_2)^2\right]^2} \\ &= x_2 \frac{(x_2)^2 - (x_1)^2}{\left[(x_1)^2 + (x_2)^2\right]^2} \end{aligned}$$

and by symmetry $(\partial_2 f)(x) = x_1 \frac{(x_1)^2 - (x_2)^2}{\left[(x_1)^2 + (x_2)^2\right]^2}$. At $x = 0$ we have

$$\begin{aligned} (\partial_1 f)(0) &\equiv \lim_{t \to 0} \frac{f(0 + t\hat{e}_1) - f(0)}{t} \\ &= \lim_{t \to 0} \frac{f\left(\begin{bmatrix} t \\ 0 \end{bmatrix}\right)}{t} \\ &= \lim_{t \to 0} \frac{\frac{t0}{t^2 + 0^2}}{t} \\ &= 0 \end{aligned}$$

and similarly by symmetry $(\partial_2 f)(0) = 0$. □

1.11. **Corollary.** *As a result we see that even though for this f the partial derivatives exist everywhere, f is not differentiable, and so it is clear that existence of partial derivatives do not necessarily imply that f is differentiable. Using Theorem 9.21 in* [1] *we see that we would need the partial derivatives to also be continuous for f to be differentiable, which, in this case, they are not (as you should verify).*

2. Hints for Solving Homework Number 3

2.1. Question 1.

- There is nothing to this question other than computing many partial derivatives. You will need to use Theorem 9.21 in [1] to conclude from the partial derivatives that your maps are indeed differentiable.

2.2. Question 2.

- Problem 9.14 in [1]. Be careful of the derivative of f at 0. Try a lucky guess and then verify that it is indeed the derivative at 0.

2.3. Question 4.

- Problem 9.10 in [1]. Thus, try a condition on U, such as convexity. Try to find a weaker condition on U.

2.4. Question 3.

- The "catch" here is that there is a product of two Banach spaces, and this defines a new Banach space with its corresponding norm, as defined on the page.

- Use the definition, together with the "guess" that $\beta'\left(\begin{bmatrix}y_1\\y_2\end{bmatrix}\right)=\left[\beta(-,y_2)\quad\beta(y_1,-)\right]$ so that $\begin{bmatrix}\tilde{y}_1\\\tilde{y}_2\end{bmatrix}\overset{\beta'\left(\begin{bmatrix}y_1\\y_2\end{bmatrix}\right)}{\mapsto}\beta(\tilde{y}_1,y_2)+\beta(y_1,\tilde{y}_2)$. Then show that this adheres to 1.1.
- Use the chain rule ([1] Theorem 9.15):

  2.1. *Claim.* Let $E\in\mathrm{Open}(X)$, Z be a Banach space, and $g:U\to Z$ where $U\in\mathrm{Open}(Y)$ such that $U\supseteq f(E)$. If f is differentiable at x and g is differentiable at $f(x)$ then the mapping $g\circ f:E\to Z$ defined by $(g\circ f)(x)\equiv g(f(x))$ is differentiable at x and $(g\circ f)'(x)=[g'(f(x))]\circ[f'(x)]$.

  Using this, $g=\beta\circ f$ and so $g'(x)=[\beta'(f(x))]\circ[f'(x)]$, where $f:X_1\times X_2\to Y_1\times Y_2$ $(x_1,x_2)\overset{f}{\mapsto}(f_1(x_1),f_2(x_2))$. Then $f'((x_1,x_2))=\left[f_1'(x_1)\quad f_2'(x_2)\right]$. Now use (a).

3. Review of Homework Number 1

- We will (hopefully) review question: 2 (partly), 5 (second part), and 3 and 4 (b) if there is time. You may find the full discussion in the solutions.

References

[1] Walter Rudin. *Principles of Mathematical Analysis (International Series in Pure and Applied Mathematics)*. McGraw-Hill Science/Engineering/Math, 1976.

- Use the definition, together with the "guess" that $\beta'\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = [\beta(-, y_2) \quad \beta(y_1, -)]$ so that $\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} \overset{\beta'\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right)}{\mapsto} \beta(\tilde{y}_1, y_2) + \beta(y_1, \tilde{y}_2)$. Then show that this adheres to 1.1.
- Use the chain rule ([1] Theorem 9.15):

2.1. *Claim.* Let $E \in \text{Open}(X)$, Z be a Banach space, and $g : U \to Z$ where $U \in \text{Open}(Y)$ such that $U \supseteq f(E)$. If f is differentiable at x and g is differentiable at $f(x)$ then the mapping $g \circ f : E \to Z$ defined by $(g \circ f)(x) \equiv g(f(x))$ is differentiable at x and $(g \circ f)'(x) = [g'(f(x))] \circ [f'(x)]$.

Using this, $g = \beta \circ f$ and so $g'(x) = [\beta'(f(x))] \circ [f'(x)]$, where $f : X_1 \times X_2 \to Y_1 \times Y_2$ $(x_1, x_2) \overset{f}{\mapsto} (f_1(x_1), f_2(x_2))$. Then $f'((x_1, x_2)) = [f_1'(x_1) \quad f_2'(x_2)]$. Now use (a).

3. Review of Homework Number 1

- We will (hopefully) review question: 2 (partly), 5 (second part), and 3 and 4 (b) if there is time. You may find the full discussion in the solutions.

References

[1] Walter Rudin. *Principles of Mathematical Analysis (International Series in Pure and Applied Mathematics)*. McGraw-Hill Science/Engineering/Math, 1976.

Problem 9.24 in Rudin

$$f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2,$$

$$(x,y) \longmapsto \Big(\underbrace{\frac{x^2-y^2}{x^2+y^2}}_{f_x}, \underbrace{\frac{xy}{x^2+y^2}}_{f_y}\Big)$$

$$\partial_x f_x = \frac{2x(x^2+y^2)-(x^2-y^2)2x}{(x^2+y^2)^2} = \frac{4xy^2}{(x^2+y^2)^2}$$

$$\partial_y f_x = \frac{-2y(x^2+y^2)-(x^2-y^2)2y}{(x^2+y^2)^2} = \frac{-4x^2y}{(x^2+y^2)^2}$$

$$\partial_x f_y = \frac{y(x^2+y^2)-xy\,2x}{(x^2+y^2)^2} = \frac{y(y^2-x^2)}{(x^2+y^2)^2}$$

$$\partial_y f_y = \frac{x(x^2+y^2)-xy\,2y}{(x^2+y^2)^2} = \frac{x(x^2-y^2)}{(x^2+y^2)^2}$$

$$\Rightarrow f'(x,y) = \begin{bmatrix} \partial_x f_x & \partial_y f_x \\ \partial_x f_y & \partial_y f_y \end{bmatrix} = \frac{1}{(x^2+y^2)^2}\begin{bmatrix} 4xy^2 & -4x^2y \\ y(y^2-x^2) & x(x^2-y^2) \end{bmatrix}$$

Observe that multiplying the first column by $-\frac{x}{y}$ gives the second column $\Rightarrow$ $\boxed{\text{rank}(f'(x,y)) \le 1}$.

But rank of a transf. is the dimension of its image.

Claim: $f_x^2 + 4f_y^2 = 1$

$\Rightarrow \text{im}(f) \subseteq \text{ellipse} \equiv 1\text{-dim space}$

More details: Rudin [1] 9.23.

$f: \mathbb{R}^2 \to \mathbb{R}$

$(x,y) \mapsto \begin{cases} xy \frac{x^2-y^2}{x^2+y^2} & \text{otherwise} \\ 0 & (x,y)=(0,0) \end{cases}$

($\partial_x \partial_y f \neq \partial_y \partial_x f$ if these derivatives are not cont. (Rudin [T] 9.41))

$$\partial_x f = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} \quad \text{if} \quad (x,y) \neq (0,0) \qquad \partial_x f = 0 \text{ if } (x,y)=0$$

$$\partial_y \partial_x f = \frac{(x-y)(x+y)(x^4+10x^2y^2+y^4)}{(x^2+y^2)^3}$$

$$\partial_y f = \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} \quad \text{if} \quad (x,y) \neq (0,0)$$

$$\partial_x \partial_y f = \frac{(x-y)(x+y)(x^4+10x^2y^2+y^4)}{(x^2+y^2)^3}$$

Problem 9.27 in Rudin:

$$f(x,y) := \begin{cases} 0 & (x,y)=(0,0) \\ \frac{xy(x^2-y^2)}{x^2+y^2} & \text{otherwise} \end{cases}$$

<u>Moral</u>: $(\partial_x \partial_y f)(0,0) \neq (\partial_y \partial_x f)(0,0)$

<u>Reason</u>: $\partial_x \partial_y f$ and $\partial_y \partial_x f$ are not continuous at $(0,0)$!

$$(\partial_x f)(x,y) = \begin{cases} \lim_{t\to 0} \frac{f(t,0)-f(0,0)}{t} = 0 & (x,y)=0 \\ \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} & (x,y) \neq 0 \end{cases}$$

$$(\partial_y \partial_x f)(x,y) = \begin{cases} \lim_{t\to 0} \frac{(\partial_x f)(0,t) - \overbrace{(\partial_x f)(0,0)}^{0}}{t} = \lim_{t\to 0} \frac{-t}{t} = -1 & (x,y)=0 \\ \frac{(x^2-y^2)(x^4+10x^2y^2+y^4)}{(x^2+y^2)^3} & (x,y) \neq 0 \end{cases}$$

$$(\partial_y f)(x,y) = \begin{cases} \lim_{t\to 0} \frac{f(0,t)-f(0,0)}{t} = 0 & (x,y)=0 \\ \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & (x,y) \neq 0 \end{cases}$$

$$(\partial_x \partial_y f)(x,y) = \begin{cases} \lim_{t\to 0} \frac{(\partial_y f)(t,0) - (\partial_y f)(0,0)}{t} = \lim_{t\to 0} \frac{t}{t} = 1 & (x,y)=0. \\ \frac{(x^2-y^2)(x^4+10x^2y^2+y^4)}{(x^2+y^2)^3} & (x,y) \neq 0 \end{cases}$$