Review

Jonathan Zhang

1 Differential Calculus in Banach Spaces and the Calculus of Variations

1.1 Differentiation

Our goal will be to generalize the notion of the 1D derivative, in the context of results like the Mean Value Theorem and Taylor's Theorem. Let us start with what we know: the standard 1D definition of the derivative is

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}, \tag{1}$$

but it is not clear how $1/h$ will generalize in higher dimensions. Indeed, it makes no sense. We can, however, define a similar concept.

Definition 1.1 (Little o). *Let X, Y be NLS's, $f : X \to Y$. If*

$$\lim_{h \to 0} \frac{\|f(h)\|_Y}{\|h\|_X} = 0,$$

we say that f is "little oh" of h and write

$$\|f(h)\|_Y = o\left(\|h\|_X\right).$$

Here, we have generalized the notion of $1/h$ to multiple dimensions by looking only at the norm of the increment and of the output. We can now formulate a notion of the derivative.

Definition 1.2 (Fréchet Derivative). *Let X, Y NLS, $U \subset X$ open, $f : U \to Y$. We say that f is Fréchet differentiable (or strongly differentiable) at a point $x \in U$ if and only if there exists a bounded linear map $A \in B(X, Y)$ such that the remainder*

$$R(x, h) := f(x+h) - f(x) - Ah$$

is $o\left(\|h\|_X\right)$. That is,

$$\lim_{h \to 0} \frac{\|R(x,h)\|_Y}{\|h\|_X} = 0.$$

If A exists, we call A the Fréchet derivative of f at x, denoted $Df(x)$. Note that by construction, $A = Df(x)$ is a linear operator on h. The definition of the Fréchet derivative is consistent with the one dimensional analogue. For $f \in C^1(\mathbb{R})$,

$$R\left(x, h\right) = f\left(x+h\right) - f\left(x\right) - f'\left(x\right) h = \left[\frac{f\left(x+h\right) - f\left(x\right)}{h} - f'\left(x\right)\right] h, \tag{2}$$

and

$$\lim_{h\to 0}\frac{|R(x,h)|}{|h|}=\lim_{h\to 0}\left[\frac{f(x+h)-f(x)}{h}-f'(x)\right]=0. \tag{3}$$

More generally, the map Df (as contrasted with $Df(x)$) can be regarded as a map $Df : U\times X \ni (x,h)\to Df(x)h\in Y$.

Proposition 1.1. *If f is Fréchet differentiable, then $Df(x)$ is unique, and f is continuous at x. Moreover, if g is Fréchet differentiable, then for every $\alpha,\beta\in\mathbb{F}$,*

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

(That is to say, the operator D taking functions $f\mapsto Df$ is linear.)

Proof. Suppose that there are two derivatives at x, $A, B\in B(X,Y)$. Define the two remainders

$$\begin{aligned}R_A(x,h)&=f(x+h)-f(x)-Ah\\R_B(x,h)&=f(x+h)-f(x)-Bh.\end{aligned} \tag{4}$$

By assumption, both $R_A(x,h)=o(\|h\|_X)$ and $R_B(x,h)=o(\|h\|_X)$. Now we show that in fact, $A=B$.

$$\begin{aligned}\|A-B\|_{B(X,Y)}&=\sup_{\|h\|=\varepsilon}\frac{\|Ah-Bh\|_Y}{\|h\|_X}\\&=\sup_{\|h\|=\varepsilon}\frac{\|Ah-Bh\|_Y}{\|h\|_X}\\&=\sup_{\|h\|=\varepsilon}\frac{\|R_A(x,h)-R_B(x,h)\|_Y}{\|h\|_X}\\&\leqslant\sup_{\|h\|=\varepsilon}\frac{\|R_A(x,h)\|_Y}{\|h\|_X}+\sup_{\|h\|=\varepsilon}\frac{\|R_B(x,h)\|_Y}{\|h\|_X}.\end{aligned} \tag{5}$$

Is this the right definition of the norm on $B(X,Y)$? The RHS can be made small as $\varepsilon\to 0$. Really? I thought the operator norm doesn't depend on this...Indeed, the argument of the supremum goes to zero as $\|h\|\to 0$ by the $o(\|h\|_X)$ assumption. Thus, $A=B=Df(x)$ is unique. To prove that f is continuous at x, we have

$$\begin{aligned}\|f(x+h)-f(x)\|_Y&=\|Df(x)h+R(x,h)\|_Y\\&\leqslant\|Df(x)h\|_Y+\|R(x,h)\|_Y\\&\leqslant\|Df(x)\|_{B(X,Y)}\|h\|_X+\|R(x,h)\|_Y.\end{aligned} \tag{6}$$

The RHS tends to 0 as $h\to 0$ in X. Since $Df(x)$ is bounded, the first term on the right tends to zero trivially as $h\to 0$. Since $R(x,h)$ is $o(\|h\|)$, it means that $R(x,h)$ goes to zero faster than h

goes to zero. This implies, however, that if $h \to 0$, then also $R(x, h) \to 0$. A somewhat obvious fact. $\square$

Having differentiability, namely *strong* differentiability, at a point x is more than enough to conclude that f is continuous at x. A piecewise linear function may be continuous, but not differentiable at the interface points. Strong differentiability allows us to conclude more:

Lemma 1.1 (Local Lipschitz Property). *If $f : U \to Y$ is differentiable at $x \in U$, then for every $\varepsilon > 0$, there exists a $\delta = \delta(x, \varepsilon) > 0$ such that for every h with $\|h\|_X \leq \delta$,*

$$\|f(x+h) - f(x)\|_Y \leqslant \left(\|Df(x)\|_{B(X,Y)} + \varepsilon\right) \|h\|_X.$$

Proof. We have

$$f(x+h) - f(x) = Df(x)\,h + R(x,h)\,. \tag{7}$$

Since the remainder is $o(\|h\|)$, then for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\frac{\|R(x,h)\|_Y}{\|h\|_X} < \varepsilon \tag{8}$$

whenever $\|h\|_X < \delta$. Definition of convergence of $\|R\|/\|h\| \to 0$. It follows then that

$$\|f(x+h) - f(x)\|_Y = \|Df(x)\,h + R(x,h)\| \leqslant \|Df(x)\|_{B(X,Y)}\|h\|_X + \varepsilon\|h\|_X. \tag{9}$$

$\square$

1.2 Fixed Points and Contractive Maps

1.3 Nonlinear Equations

1.4 Higher Derivatives

1.5 Extrema

1.6 Euler-Lagrange Equations

1.7 Constrained Extrema and Lagrange Multipliers