Lecture 25 : Frechet Derivative & Linearization of Maps ①

Our definition for $f : \mathbb{R} \to \mathbb{R}$ where $f'(a)$ exists iff $\lim_{h\to 0}\left(\frac{f(a+h)-f(a)}{h}\right) \in \mathbb{R}$ fails to generalize directly for $F : \mathbb{R}^n \to \mathbb{R}^m$. There are a couple ways to anticipate this,

① we can't (usually) divide by a vector
② to capture the analog of the tangent line for $F : \mathbb{R}^n \to \mathbb{R}^m$ we need much more data.

If $n = 1, m > 1$ it's space curve so need a vector to capture derivative ok... in <u>this</u> case $\frac{d\vec{r}}{dt}\Big|_{t_0} = \lim_{h\to 0}\left(\frac{\vec{r}(t_0+h)-\vec{r}(t_0)}{h}\right)$

But, $n > 1$, $m = 1$ need ∇f to capture tangent space for $f : \mathbb{R}^2 \to \mathbb{R}$ for example...

Motivations aside, this is how to define the <u>derivative</u> of $F : \mathbb{R}^n \to \mathbb{R}^m$ in general,

Defn/ If $F : \mathbb{R}^n \to \mathbb{R}^m$ and $\exists\, L : \mathbb{R}^n \to \mathbb{R}^m$ linear transformation such that the following holds:

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

$\underbrace{\lim_{h\to 0}(F(a+h)) = F(a)}_{\text{continuity of } F \text{ at } a.}$

then we say F is <u>differentiable</u> at $a \in \mathbb{R}^n$ and write $dF_a(h) = L(h)$. We say dF_a is the <u>differential</u> of F at a and define $J_F(a) = [dF_a]$ to be the <u>Jacobian</u> <u>Matrix</u> of F at a.

Th^m/ If $F : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a then $J_F(a) = \left[\frac{\partial F}{\partial x_1}(a) \,\middle|\, \frac{\partial F}{\partial x_2}(a) \,\middle|\, \cdots \,\middle|\, \frac{\partial F}{\partial x_n}(a) \right]$.

That is, omitting a, $(J_F)_{ij} = \frac{\partial F_i}{\partial x_j} = \partial_j F_i$

Proof: Suppose F differentiable at a. Then

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

The existence of the above multivariate limit implies all path limits likewise converge to 0. Consider $t \mapsto te_i$ for $t \to 0$ then $h = te_i$ provides

$$\lim_{t \to 0} \left(\frac{F(a+te_i) - F(a) - dF_a(te_i)}{|t|} \right) = 0$$

For $t \to 0^+$ obtain

$$\lim_{t \to 0^+} \left(\frac{F(a+te_i) - F(a)}{t} - \frac{t\, dF_a(e_i)}{t} \right) = 0$$

Hence $\underbrace{\lim_{t \to 0^+} \left(\frac{F(a+te_i) - F(a)}{t} \right)}_{\star} = dF_a(e_i) = \mathrm{col}_i [dF_a]$
$= \mathrm{col}_i (J_F(a))$

and the same holds for $t \to 0^-$. Notice $\star$ (adding $t \to 0^-$ as well) is precisely $\frac{\partial F}{\partial x_i}(a)$ thus $\mathrm{col}_i([dF_a]) = \frac{\partial F}{\partial x_i}(a)$. //

Th^m/ If $F : \mathbb{R}^n \to \mathbb{R}^m$ has all partial derivatives $\partial_j F_i$ exist, and continuously so, then F is differentiable at points where J_F exists.

(continuously differentiable at a $\Rightarrow$ differentiable at a)

Proof: somewhat involved, also quite beautiful. See Edwards Th^m 2.5, pg. 72

Remark: the assignment $a \mapsto dF_a$ is a map from $\mathbb{R}^n$ to linear maps from $\mathbb{R}^n \to \mathbb{R}^m$ that is not something we have explained the meaning of continuity for... but it can be described via the "operator norm" and when you sort through all it then you'll be pleased to learn $\underbrace{\text{continuously diff at } a}_{\text{statement about continuity of partial derivatives.}} \Rightarrow \underbrace{a \mapsto dF_a \text{ continuous.}}_{\text{above our pay grade here.... but it's in Edwards, Prop 2.4 pg. 176.}}$

PRACTICAL INTERPRETATION

If $F : \mathbb{R}^n \to \mathbb{R}^m$ is diff. at a then

$$F(x) \cong \underbrace{F(a) + J_F(a)(x-a)}_{\text{Linearization of } F \text{ at } a.}$$

Or,

$$F(a+h) \cong \underbrace{F(a) + J_F(a)h}_{\text{Linearization of } F \text{ at } a.\ \text{Def}^{\text{n}}\ L_F^a(h)}$$

So, we use Jacobian matrix to find best linear approximation to generally "curved" F.

Example 1: $F(x,y) = \sqrt{1+x^2+y^2}$

$F(2,2) = \sqrt{1+4+4} = \sqrt{9} = 3$

$J_F = \left[\frac{\partial F}{\partial x} \middle| \frac{\partial F}{\partial y}\right] = \left[\frac{x}{\sqrt{1+x^2+y^2}} \middle| \frac{y}{\sqrt{1+x^2+y^2}}\right]$

$J_F(2,2) = [2/3 \,|\, 2/3]$

$F(2+h, 2+k) \approx 3 + [2/3 \,|\, 2/3]\begin{bmatrix} h \\ k \end{bmatrix} = \boxed{3 + \overbrace{\frac{2h+2k}{3}}^{L_F^{(2,2)}(h,k)}}$

Then $L_F^{(2,2)}(1,0) = 3 + \frac{2}{3} = \underset{3.66}{\frac{11}{3}} \cong \sqrt{1+9+4} = \underset{3.74}{\sqrt{14}}$

vs.

Example 1 continued

$$L_F^{(2,2)}(x,y) = F(2,2) + J_F(2,2)\begin{bmatrix} x-2 \\ y-2 \end{bmatrix}$$

$$= 3 + \left[\tfrac{2}{3} \,\middle|\, \tfrac{2}{3}\right]\begin{bmatrix} x-2 \\ y-2 \end{bmatrix}$$

$$= 3 + \frac{2}{3}(x-2) + \frac{2}{3}(y-2).$$

Remark: I shouldn't use same $L_F^{(1,1)}$ for both the map with (x,y) inputs and $(a+h, a+h)$ inputs.

Example 2:

$$F(x,y,z) = (x+y+z,\ x^2+z^2,\ y^2-z^2)$$

$$J_F = \left[\frac{\partial F}{\partial x} \,\middle|\, \frac{\partial F}{\partial y} \,\middle|\, \frac{\partial F}{\partial z}\right] = \begin{bmatrix} 1 & 1 & 1 \\ 2x & 0 & 2z \\ 0 & 2y & -2z \end{bmatrix}$$

Let's calculate the linearization at $(1,2,3)$

$$L_F^{(1,2,3)}(x,y,z) = F(1,2,3) + J_F(1,2,3)\begin{bmatrix} x-1 \\ y-2 \\ z-3 \end{bmatrix}$$

$$L_F^{(1,2,3)}(x,y,z) = \begin{bmatrix} 6 \\ 10 \\ -5 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 6 \\ 0 & 4 & -6 \end{bmatrix}\begin{bmatrix} x-1 \\ y-2 \\ z-3 \end{bmatrix}$$

$$L_F^{(1,2,3)}(x,y,z) = \big(6 + 1\cdot(x-1) + (y-2) + z-3,\ 2(x-1) + 6(z-3),\ 4(y-2) - 6(z-3)\big)$$

$$L_F^{(1,2,3)}(x,y,z) = (x+y+z,\ 2x+6z-20,\ 4y-6z+10).$$

Remark: I've given the def$^{=}$ for maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, but this definition easily extends to NORMED LINEAR SPACES, like $\mathbb{R}^{n\times n}$ with $\|A\| = \text{trace}(A^TA)$ etc. See my MATH 332 notes for more about such calculus...