Lecture IX: Differentiability of Functions of Several Variables

Definition: A function $F: \mathcal{U} \to \mathbb{R}^m$, $\mathcal{U}$ open in $\mathbb{R}^n$, is <u>differentiable</u> <u>at</u> $\vec{x}_0 \in \mathcal{U}$ if $\exists$ a linear function $L: \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\vec{h} \to 0} \frac{1}{\|\vec{h}\|_n} \| F(\vec{x}_0 + \vec{h}) - F(\vec{x}_0) - L(\vec{h}) \|_m = 0.$$

Here $\| \ \|_n$ refers to the standard Euclidean "norm" on $\mathbb{R}^n$, $\|(a_1, \dots, a_n)\| = (\sum_1^n a_j^2)^{1/2}$. However, the definition would be the same if these standard norms were replaced by any other vector space norms on $\mathbb{R}^n$ and $\mathbb{R}^m$, these being necessarily "equivalent" (uniformly comparable: cf. supplementary notes on norm equivalence on $\mathbb{R}^n$). Notation $L = F'|_{x_0}$ or $L = F'(x_0)$

If $n = m = 1$, this definition of differentiability is the same as the usual one-variable definition, $L(\vec{h})$ in the $n=m=1$ case being (multiplication by) $f'(x_0)$.

It is often convenient to express the differentiability condition in the "little o" notation:
F is differentiable at $\vec{x}_0$ if, for some linear function L,

$$F(\vec{x}_0 + \vec{h}) = F(\vec{x}_0) + L\vec{h} + o(\|\vec{h}\|).$$

Here $o(\|\vec{h}\|)$ means a term small compared to $\|\vec{h}\|$ when $\|\vec{h}\|$ is small, i.e. $\lim_{\vec{h}\to 0} \frac{1}{\|\vec{h}\|}$ "$o(\|h\|)$" $= 0$

The most fundamental estimate involving differentiation of functions of one-variable comes from the mean value property: Since $f(b)-f(a) = f'(\lambda)(b-a)$ for some $\lambda \in (a,b)$, $|f(b)-f(a)| \le M(b-a)$ if $|f'| \le M$ on (a,b). The mean value equality does not have an exact analogue in the several variable case: a curve in $\mathbb{R}^3$ can go from $\vec{a}$ to $\vec{b}$ without ever having its velocity parallel to $\vec{b}-\vec{a}$ for example. But the associated inequality does have an analogue.

To state this, recall the concept of $\|L\|$ for L a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$:
$\|L\|$ ("the operator norm of L") is by definition
$= \{\sup \|L(\vec{v})\|_m : \vec{v} \in \mathbb{R}^n \ni \|v\| \le 1\}$.
Since L is continuous and $\{\vec{v} : \|v\| \le 1\}$, the sup is actually a max: it is attained at some point. (Note: $\|L\| = \{\max \|L(\vec{v})\|_m : \vec{v} \in \mathbb{R}^n, \|\vec{v}\| = 1\}$ because L is linear).

Mean Value Inequality for Differentiable Functions in the Multi-Variable Case:
Suppose $F: U \to \mathbb{R}^m$ is a differentiable function defined on an open subset U of $\mathbb{R}^n$. Suppose $\vec{a}, \vec{b} \in U$ and that the line segment $\mathcal{L}_{a,b} = \{\lambda\vec{a} + (1-\lambda)\vec{b} : \lambda \in [0,1]\} \subset U$. Then

$$\|F(\vec{b}) - F(\vec{a})\| \le (\sup \|F'\|)(\|b-a\|)$$

Here the $\sup \|F'\|$ need only be taken over points of the line segment $\mathcal{L}_{\vec{a},\vec{b}}$.

(Note: Since we are not assuming F' is continuous, the sup might in principle be $+\infty$, in which case, one obtains no information of course).

Proof of the Theorem: It is enough to show that $< \vec{u}, F(\vec{b}) - F(\vec{a}) >$

$\leq (\sup \|F'\|) \cdot (\|\vec{b} - \vec{a}\|)$

for each $\vec{u} \in \mathbb{R}^m$ with $\|\vec{u}\| = 1$. (If $F(\vec{b}) - F(\vec{a}) = 0$, we are done anyway. If $F(\vec{b}) - F(\vec{a}) \neq 0$, set $\vec{u} = (F(\vec{b}) - F(\vec{a}))/\|F(\vec{b}) - F(\vec{a})\|$)

For this, let

$$f(\lambda) = < \vec{u}, F(\lambda \vec{a} + (1-\lambda)\vec{b})>$$

Then f is a differentiable function on $(0,1)$ and $f'(\lambda) = < \vec{u}\ L(\vec{a} - \vec{b})>$
where $L = F'$ at $\lambda \vec{a} + (1-\lambda)\vec{b}$.
[This comes from the Chain Rule: see later].
So $|f'(\lambda)| \leq (\sup \|F'\|)(\|\vec{b} - \vec{a}\|)$
But $f(1) = F(\vec{a})$ and $f(0) = F(\vec{b})$
So

$$\|f(1) - f(0)\| = \|F(\vec{a}) - F(\vec{b})\|$$

$$= \|F(\vec{b}) - F(\vec{a})\| \leq (\sup |f'|)(1-0)$$

$$\leq (\sup \|F'\|)(\|\vec{b} - \vec{a}\|) \quad \square.$$

The MV inequality has an intuitive meaning: Think of the (linear) curve $\gamma(t) = \lambda\vec{a} + (1-\lambda)\vec{b}$. It has velocity vector $-\vec{b} + \vec{a}$ and hence has constant speed $\|\vec{b} - \vec{a}\|$. The speed of its image under F, that is, $F(\gamma(t))$ is $\leq (\sup\|F'\|)(\|\vec{b} - \vec{a}\|)$ since the velocity vector of $F(\gamma(t))$ is $F'(\vec{b} - \vec{a})$. So in "time" 1, as λ changes from 0 to 1, the curve $F(\gamma(t))$ cannot move more than (time elapsed)(max. speed) $\leq (1)(\sup\|F'\|)(\|\vec{b} - \vec{a}\|)$.

This estimate is central to understanding how functions of several variables behave. In particular, it will play a fundamental role in the proof of the Inverse Function Theorem.

Before beginning that subject, we list some important but elementary matters, which are mostly left as exercises or given only outlines of proof, with details left to you:

<u>Exercises</u> 1. $F: U \to \mathbb{R}^m$ is differentiable at $x_0 \in U$ $\iff$ each of the "components" of F are differentiable at x_0 [If $F(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x}))$ where f_j is $\mathbb{R}$-valued, then each f_j is differentiable at $\vec{x}_0$]

2. In the component notation ~~general~~ just introduced, the matrix of F' at x_0, if F is differentiable at x_0, is $\left(\partial f_j / \partial x_i\right)\Big|_{\vec{x}_0}$ $\quad i = 1, \dots n$, $\quad j = 1, \dots m$ $\qquad \vec{x} = (x_1, \dots, x_n)$

(and differentiability of F at $x_0 \Rightarrow$ these partial derivative exist)

3. If $F: \mathcal{U} \to \mathbb{R}$, $\mathcal{U} \subset \mathbb{R}^n$, and if the partial derivatives $\partial F/\partial x_j$, $j = 1, 2 \ldots n$ exist at each point of $\mathcal{U}$ and are continuous on $\mathcal{U}$, then F is differentiable at each point $x_0 \in \mathcal{U}$ with $F'|_{x_0}$ = the linear transformation

$$F'|_{(x_0)}(h_1, \ldots h_n) = \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\bigg|_{x_0} h_j$$

[Use one-var. MVTheorem for proof]

"Chain Rule" 4. If $F: \mathcal{U} \to \mathbb{R}^m$ and $G: V \to \mathbb{R}^k$, $V^{\text{open}} \subset \mathbb{R}^m$, $\mathcal{U}^{\text{open}} \subset \mathbb{R}^n$, if F is differentiable at $x_0 \in \mathcal{U}$, if $F(x_0) \in V$, and G is differentiable at $F(x_0)$, then $G(F(\cdot))$ is differentiable at x_0 and its derivative is the linear function that takes $\vec{h} \in \mathbb{R}^n$ to $G'|_{F(x_0)}(F'|_{x_0}(\vec{h}))$.

Proof sketch: $F(\vec{x}_0 + \vec{h}) = F(\vec{x}_0) + L_1\vec{h} + o(\vec{h})$ where $L_1 = F'|_{x_0}$. Since $L_1\vec{h} + o(\vec{h}) \to \vec{0}$ as $\vec{h} \to 0$, with $L_2 = G'|_{F(x_0)}$:

$$G(F(x_0+h)) = G(F(x_0)) + L_2(L_1\vec{h} + o(\vec{h})) + o(L_1\vec{h} + o(\vec{h}))$$

Now $\|L_2 v\| \leq \|L_2\|\|v\|$ so $L_2(o(\vec{h}))$ is $o(\vec{h})$ &

$$L_2(L_1\vec{h} + o(\vec{h})) = L_2(L_1\vec{h}) + o(\vec{h})$$

Also $o(L_1\vec{h} + o(\vec{h}))$ is $o(\vec{h})$. $\square$

Examples and exercises on chain rule

Polar coordinates: (r,θ) $\quad x = r\cos\theta,\ y = r\sin\theta$

$r = \sqrt{x^2+y^2} \quad \theta =$ (e.g.) $\arctan(y/x)$

$$\frac{\partial x}{\partial r} = \cos\theta \quad \frac{\partial x}{\partial \theta} = -r\sin\theta \quad \frac{\partial y}{\partial r} = \sin\theta \quad \frac{\partial y}{\partial \theta} = r\cos\theta$$

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} \quad \frac{\partial \theta}{\partial x} = \frac{-y/x^2}{1+\left(\frac{y}{x}\right)^2} = \frac{-y}{x^2+y^2} \quad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} \quad \frac{\partial \theta}{\partial y} = \frac{x}{x^2+y^2}$$

Should have

$$1 = \frac{\partial r}{\partial r} = \frac{\partial r}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial r}{\partial y}\frac{\partial y}{\partial r} = \frac{x}{\sqrt{x^2+y^2}}\cdot\cos\theta + \frac{y}{\sqrt{x^2+y^2}}\sin\theta$$

$$= \frac{x}{r}\cdot\frac{x}{r} + \frac{y}{r}\cdot\frac{y}{r} = \frac{x^2+y^2}{r^2} = 1 \ \checkmark$$

Exercise: Check similarly $\frac{\partial\theta}{\partial r} = 0 \quad \frac{\partial r}{\partial\theta} = 0 \quad \frac{\partial\theta}{\partial\theta} = 1$

and $1 = \frac{\partial x}{\partial x}$, $1 = \frac{\partial y}{\partial y}$, $0 = \frac{\partial x}{\partial y}$, $0 = \frac{\partial y}{\partial x}$.

Exercise: $$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial\theta^2}$$