MATH 395 Notes

Faye Jackson

September 28, 2020

Continuously Differentiable Functions

We saw that if $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at x then $Df(x)$ is given by the partial derivatives as:

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

The converse statement that if the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist for each $1 \le i \le m$ and $1 \le j \le n$ then Df exists is FALSE. However we have a slightly stronger condition that works!

Theorem. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open. Suppose that all partial derivatives $\frac{\partial f_i}{\partial x_j}$ $(1 \le i \le m, 1 \le j \le n)$ exist in some neighborhood of $x \in A$, and they are continuous at x.*

Then f is differentiable at x. In particular if all partial derivatives exist and are continuous throughout A then f is differentiable in A. Such a function is called a C^1 function.

Remark. This theorem allows us to recognize "most" differentiable functions that we meet in practice just by checking that the partials are continuous.

Proof. Since f is differentiable at x if and only if each of its component functions are differentiable at x, we may assume without loss of generality that $m = 1$.

Let $r > 0$ be such that $B(x, r) \subseteq A$ and the partials are defined and continuous on $B(x,r)$. Then let $h \in \mathbb{R}^n$ such that $\|h\| < r$. Let $h = (h_1, \ldots, h_n)$. Set:

$$\begin{aligned} p_0 &:= x \\ p_k &:= p_{k-1} + h_k e_k \end{aligned}$$

And so $p_n = x + h$. So then we have:

$$f(x+h) - f(x) = \sum_{j=1}^{n} f(p_j) - f(p_{j-1})$$

Now we know that:

$$f(p_j) - f(p_{j-1}) = f(p_{j-1} + h_j e_j) - f(p_j)$$

Define $\phi_j(s) := f(p_{j-1} + s e_j)$ where ϕ is defined on some neighborhood of 0 in $\mathbb{R}$. Since ϕ_j is differentiable on an open interval containing $[0, h_j]$ with derivative $\frac{\partial f}{\partial x_j}$, we know that ϕ_j is continuous on $[0, h_j]$ and differentiable on $(0, h_j)$. Therefore by the mean value theorem we know that for some $c_j^h \in (0, h_j)$ that:

$$\begin{aligned} \phi_j(h_j) - \phi(0) &= h_j \cdot \phi'(c_j^h) \\ f(p_j) - f(p_{j-1}) &= h_j \cdot \frac{\partial f}{\partial x_j}(p_{j-1} + c_j^h e_j) \\ &= h_j \cdot \frac{\partial f}{\partial x_j}(q_j) \end{aligned}$$

Where q_j is some point in $B(x, \|h\|)$. Therefore:

$$f(x+h) - f(x) = \sum_{i=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(q_j)$$

For some $q_1, \ldots, q_n$ in the ball of radius $\|h\|$ centered at x. Therefore:

$$f(x+h) - f(x) - \sum_{i=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(x) = \sum_{j=1}^{n} h_j \cdot \left[\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x)\right]$$

This implies that:

$$\begin{aligned}
\frac{\left\|f(x+h) - f(x) - \sum_{i=1}^n h_j \cdot \frac{\partial f}{\partial x_j}(x)\right\|}{\|h\|} &= \frac{\left\|\sum_{j=1}^n h_j \cdot \left[\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x)\right]\right\|}{\|h\|} \\
&\overset{\triangle}{\leq} \sum_{j=1}^n \frac{\|h_j\|}{\|h\|} \cdot \left\|\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x)\right\| \\
&\leq \sum_{j=1}^n \left\|\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x)\right\| \\
&\leq n \cdot \sup_{\substack{q \in B(x, \|h\|) \\ 1 \leq j \leq n}} \left\|\frac{\partial f}{\partial x_j}(q) - \frac{\partial f}{\partial x_j}(x)\right\|
\end{aligned}$$

But this goes to 0 as $\|h\| \to 0$ since $\frac{\partial f}{\partial x_j}$ are assumed to be continuous at x. Note then that we win! The function:

$$T(h) = \sum_{i=1}^n h_j \cdot \frac{\partial f}{\partial x_j}(x)$$

is a linear function, and so f is differentiable at x, and of course Df is just the vector $\left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)(x)$. Amazing!!!

Higher Order Derivatives

Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open. The component functions are $f_i : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Since $\frac{\partial f_i}{\partial x_j}$ is itself a function from $A \to \mathbb{R}$ we can take higher order partial derivatives. For instance, if $f_i \in C^1$ then $\frac{\partial f_i}{\partial x_j}$ is defined and continuous, so we can consider if the following exists:

$$\frac{\partial^2 f_i}{\partial x_k \partial x_j} := \frac{\partial}{\partial x_k} \frac{\partial f_i}{\partial x_j}$$

This is called a second-order partial derivative. Similarly one can define partial derivatives of higher order inductively.

Definition. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is of class $\underline{C^r}$ for $r \in \mathbb{N}_0$ provided that all the partial derivatives of its component functions up to order r exist and are continuous.*

We say that f is $\underline{C^\infty}$ provided that $f \in C^r$ for all $r \in \mathbb{N}_0$.

Exercise. *Show that* $f \in C^r$ *if and only if* $\frac{\partial f}{\partial x_j} \in C^{r-1}$ *for each* $1 \leq j \leq n$.

Multi-Index Notation

Definition. *A multi-index is an* n*-tuple* $\alpha = (\alpha_1, \ldots, \alpha_n)$ *such that each* $\alpha_i \in \mathbb{N}_0$. *If* α *is a multi-index then we define:*

- *The order of* α *as* $|\alpha| := \alpha_1 + \cdots \alpha_n$. *And the*
- *The factorial* $\alpha! := \alpha_1!\alpha_2! \cdots \alpha_n!$
- *For* $x \in \mathbb{R}^n$ *we define* $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$.
- *For* $f_i : \mathbb{R}^n \to \mathbb{R}$ *we define* $\partial^\alpha f := \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} f$.

Example. For $f : \mathbb{R}^2 \to \mathbb{R}$ we see:

$$\partial^{(1,2)} f = \frac{\partial^3 f}{\partial x_1 \partial x_2^2}$$

But wait, then what about $\frac{\partial^2 f}{\partial x_2 \partial x_1}$? Does it have a multi-index notation?

Theorem. *Let* $A \subseteq \mathbb{R}^n$ *be open and let* $f : A \to \mathbb{R}$ *be a function of class* C^2. *Then for each* $x \in A$ *we have:*

$$\frac{\partial^2}{\partial x_k \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x)$$

Corrolary. *If* $f : A \to \mathbb{R}$ *is of class* C^r *then for any* $2 \leq m \leq r$ *then:*

$$\frac{\partial^m f}{\partial x_{j_1} \cdots \partial x_{j_m}} = \frac{\partial^m f}{\partial x_{\widetilde{j}_1} \cdots \partial x_{\widetilde{j}_m}}$$

for any permutation $\widetilde{j}_1, \ldots, \widetilde{j}_m$ *of* $j_1, \ldots, j_m$. *In particular we can always rearrange* $j_1, \ldots, j_m$ *such that* $\widetilde{j}_1 \leq \cdots \leq \widetilde{j}_m$ *and in that case there is a multi-index notation:*

$$\frac{\partial^m f}{\partial x_{j_1} \cdots \partial x_{j_m}} = \frac{\partial^m f}{\partial x_{\widetilde{j}_1} \cdots \partial x_{\widetilde{j}_m}} = \partial^\alpha f$$

For some multi-index α. *Therefore any partial derivative up to order* r *can be written in multi-index notation as* $\partial^\alpha f$ *for some multi-index with order less than or equal to* r.

Exercise. *Deduce the corollary from the theorem using induction.*

Proof of Theorem. We start with some reductions. Since one computes $\frac{\partial f}{\partial x_i \partial x_j}$ and $\frac{\partial f}{\partial x_j \partial x_i}$ by keeping all other coordinates x_k for $k \neq i, j$ constant, we can assume without loss of generality that $n = 2$, and that $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$.