CALCULUS ON MANIFOLDS

JACK CERONI

Contents

Date: October 2021.

1. Introduction

2. Chapter 2

2.1. Notes.

Remark 1 (Motivating the Chain Rule)**.** *The chain rule is a very natural statement. More specifically, it states that given a the composition of two functions, $g \circ f$, we will have:*

$$[D(g \circ f)](a) = [Dg](f(a)) \circ [Df](a)$$

which amounts to the statement that a linear approximation near the point a of the function $g \circ f$ is the same as taking a linear approximation of f near a, and then mapping the points given from this linear approximation using the linear approximation of g near $f(a)$.

Theorem 1 (General Chain Rule)**.** *Given functions f and g, such that f is differentiable at a, and g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a, with:*

$$[D(g \circ f)](a) = [Dg](f(a)) \circ [Df](a) \tag{1}$$

Proof. Our goal is to show that in some neighbourhood around a, we have:

$$(g \circ f)(a+h) = (g \circ f)(a) + ([Dg](f(a)) \circ [Df](a))\, h + o(h)$$

where $o(h)$ is small. We first note that in a neighbourhood around $f(a)$, and a neighbourhood around a, we have:

$$g(f(a)+h) = g(f(a)) + [Dg](f(a))h + q(h) \qquad f(a+h) = f(a) + [Df](a)h + p(h)$$

where $q(h)$ and $p(h)$ are small. It follows that in this neighbourhood, we have:

$$g(f(a+h)) = g\left(f(a) + [Df](a)h + p(h)\right)$$

Since $[Df](a)h + p(h) \to 0$ as $h \to 0$, we can choose h small enough such that $f(a) + [Df](a)h + p(h)$ is in the neighbourhood of $f(a)$ for which the previous statement holds. In this neighbourhood around a, we have:

$$(g \circ f)(a+h) = (g \circ f)(a) + [Dg](f(a))\left([Df](a)h + p(h)\right) + q\left([Df](a)h + p(h)\right)$$

$$\Rightarrow (g \circ f)(a+h) - (g \circ f)(a) - ([Dg](f(a)) \circ [Df](a))\, h = [Dg](f(a))p(h) + q\left([Df](a)h + p(h)\right)$$

so all that is left to do is to show the right-hand side of the above equation is small. Indeed, we have $|[Dg](f(a))p(h)| \leq M|p(h)|$, for some M. Thus:

$$\frac{|[Dg](f(a))p(h)|}{|h|} \leq \frac{M|p(h)|}{|h|}$$

Clearly, the right-hand side of the inequality goes to 0 as $h \to 0$, so the left-hand side does as well. Finally, we prove the final function is small.

Note that $|[Df](a)h| \leq M|h|$ for some M, and since $p(h)/|h| \to 0$, we can choose $|h| < r_1$ such that $|p(h)| < |h|$. Thus, for $|h| < r_1$, we have $|[Df](a)h + p(h)| \leq (M+1)|h|$.

Given some $\epsilon > 0$, we choose $|t| < \delta$ such that $\frac{|q(t)|}{|t|} < \frac{\epsilon}{M+1}$. We also choose r_2 such that for $|h| < r_2$, we have $|[Df](a)h + p(h)| < \delta$. We then note that for $|h| < r = \min\{r_1, r_2\}$, we have:

$$\frac{q\left([Df](a)h + p(h)\right)}{|h|} \leq (M+1)\frac{q\left([Df](a)h + p(h)\right)}{|[Df](a)h + p(h)|} < (M+1)\frac{\epsilon}{M+1} = \epsilon$$

Thus, this function is small as well. □

Theorem 2. *If f is differentiable, then all of the partial derivatives exist and the matrix of $[Df](a)$ with respect to the standard basis is precisely the matrix of partial derivatives.*

Proof. For some a, we can define a function of the form $h(x) = (a_1, \ ..., \ a_{k-1}, x, a_{k+1}, \ ..., \ a_n)$ We then note that the function $f \circ h$ is differentiable, as it is the composition of two differentiable functions. It is also easy to see that:

$$(f \circ h) = (f_1 \circ h, \ ..., \ f_m \circ h)$$

where each $f_j \circ h$ is differentiable. It is easy to see that $(f_j \circ h)'(a_k) = D_k f_j(a)$, from the definition of the partial derivative, so the partial derivatives exist. In addition, we have that:

$$D_k f_j(a) = [D(f_j \circ h)](a_k) = [Df_j](a) \circ e_k$$

which is the k-entry of $[Df_j](a)$. But since $[Df_j](a)$ is the j-th row of the matrix for $[Df](a)$, it follows that $D_k f_j(a)$ is the entry at the k-column and j-th row. □

2.2. **Problems and Solutions.**