Chapter 2: Differentiation

Notes on Spivak by Patch Kessler

March 18, 2019

These notes and observations document my progression through Michael Spivak's Calculus on Manifolds[1]. The idea is to document my learning in beautiful Tufte-LaTeX style documents. Spivak's text is black, while all of my own writing is blue.

[1] Michael Spivak. *Calculus on Manifolds*. Benjamin Cummings, 1965. ISBN 0846590219

Basic Definitions

Recall that a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is differentiable at $a \in \mathbb{R}$ if there is a number $f'(a)$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a). \quad (1)$$

This equation certainly makes no sense in the general case of a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, but can be reformulated in a way that does. If $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ is the linear transformation defined by $\lambda(h) = f'(a) \cdot h$, then equation (1) is equivalent to

$$\lim_{h \to 0} \frac{f(a+h) - (f(a) + \lambda(h))}{h} = 0. \quad (2)$$

Equation (1) is often interpreted as saying that $f(a) + \lambda$ is a good approximation to f at a (see Problem 2-9). Henceforth we focus our attention on the linear transformation λ and reformulate the definition of differentiability as follows.

A function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is differentiable at $a \in \mathbb{R}$ if there is a linear transformation $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ such that

$$\lim_{h \to 0} \frac{f(a+h) - (f(a) + \lambda(h))}{h} = 0.$$

In this form the definition has a simple generalization to higher dimenstions:

A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is **differentiable** at $a \in \mathbb{R}^n$ if there is a linear transformation $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{|f(a+h) - (f(a) + \lambda(h))|}{|h|} = 0.$$

Note that h is a point of $\mathbb{R}^n$ and $f(a+h) - (f(a) + \lambda(h))$ a point of $\mathbb{R}^m$, so the norm signs are essential. The linear transformation λ is denoted $Df(a)$ and called the **derivative** of f at a. The justification for the phrase "*the* linear transformation λ" is

2-1 Theorem. If $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, there is a *unique* linear transformation $\lambda:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{|f(a+h) - (f(a) + \lambda(h))|}{|h|} = 0.$$

2-1 Proof. Suppose $\mu:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfies

$$\lim_{h\to 0} \frac{|f(a+h) - (f(a) + \mu(h))|}{|h|} = 0.$$

If $d(h) = f(a+h) - f(a)$, then

$$\begin{aligned} \lim_{h\to 0} \frac{|\lambda(h) - \mu(h)|}{|h|} &= \lim_{h\to 0} \frac{|\lambda(h) - d(h) + d(h) - \mu(h)|}{|h|} \\ &\leq \lim_{h\to 0} \frac{|\lambda(h) - d(h)|}{|h|} + \lim_{h\to 0} \frac{|d(h) - \mu(h)|}{|h|} \\ &= 0. \end{aligned}$$

If $x \in \mathbb{R}^n$, then $tx \to 0$ as $t \to 0$. Hence for $x \neq 0$ we have

$$0 = \lim_{t\to 0} \frac{|\lambda(tx) - \mu(tx)|}{|tx|} = \frac{|\lambda(x) - \mu(x)|}{|x|}.$$

Therefore $\lambda(x) = \mu(x)$. $\square$

We shall later discover a simple way of finding $Df(a)$. For the moment let us consider the function $f:\mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $f(x,y) = \sin x$. Then $Df(a,b) = \lambda$ satisfies $\lambda(x,y) = (\cos a)\cdot x$. To prove this, note that

$$\lim_{(h,k)\to 0} \frac{|f(a+h,b+k) - (f(a,b) + \lambda(h,k))|}{|(h,k)|} = \lim_{(h,k)\to 0} \frac{|\sin(a+h) - (\sin a + (\cos a)\cdot h)|}{|(h,k)|}.$$

Since $\sin'(a) = \cos a$, we have

$$\lim_{h\to 0} \frac{|\sin(a+h) - (\sin a + (\cos a)\cdot h)|}{|h|} = 0.$$

Since $|(h,k)| \geq |h|$, it is also true that

$$\lim_{h\to 0} \frac{|\sin(a+h) - (\sin a + (\cos a)\cdot h)|}{|(h,k)|} = 0.$$

It is often convenient to consider the matrix of $Df(a):\mathbb{R}^n \longrightarrow \mathbb{R}^m$ with respect to the usual bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. This $m \times n$ matrix is called the **Jacobian matrix** of f at a, and denoted $f'(a)$. If $f(x,y) = \sin x$, then $f'(a,b) = (\cos a, 0)$. If $f:\mathbb{R} \longrightarrow \mathbb{R}$, then $f'(a)$ is a 1×1 matrix whose single enbtry is the number which is denoted $f'(a)$ in elementray calculus.

The definition of $Df(a)$ could be made if f were defined only in some open set containing a. Considering only functions defined on $\mathbb{R}^n$ streamlines the statement of theorems and produces no real loss of generality. It is convenient to define a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ to be differentiable on A if f is differentiable at a for each $a \in A$. If $f : A \longrightarrow \mathbb{R}^m$, then f is called differentiable if f can be estended to a differentiable function on some open set containing A.

Problems.

2-1.* Prove that if $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then it is continuous at a. Hint: Use Problem 1-10.

If f is differentiable at a, then $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ exists such that

$$\lim_{h \to 0} \frac{|f(a+h) - (f(a) + Th)|}{|h|} = 0.$$

Once $|h| < 1$, we have

$$\frac{|f(a+h) - (f(a) + Th)|}{|h|} > |f(a+h) - (f(a) + Th)|.$$

Let $A = f(a+h) - f(a)$ and $B = Th$. From the above, we know that $|A - B|$ gets arbitrarily small with h, and so $A \to B$ as $h \to 0$. We also know that $B \to 0$ as $h \to 0$. This follows from Problem 1-10 which guarantees the existence of a number M such that $|B| = |Th| < M|h|$. If $A \to B$ and $B \to 0$, then $A \to 0$. It follows that f is continuous at a.

2-2. A function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is **independent of the second variable** if for each $x \in \mathbb{R}$ we have $f(x, y_1) = f(x, y_2)$ for all $y_1, y_2 \in \mathbb{R}$. Show that f is independent of the the second variable if and only if there is a function $g : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f(x, y) = g(x)$. What is $f'(a, b)$ in terms of g'?

Suppose f is independent of the second variable. Then define $g(x) = f(x, 0)$. Note that $g : \mathbb{R} \longrightarrow \mathbb{R}$ and that $f(x, y) = f(x, 0) = g(x)$.

In the other direction, let $g : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f(x, y) = g(x)$. Note that $f(x, y_1) = g(x)$ and that $f(x, y_2) = g(x)$. Thus $f(x, y_1) = f(x, y_2)$, and so f is independent of the second variable.

The derivative $f'(a, b)$ is the $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ for which

$$\lim_{h_1, h_2 \to 0} \frac{|f(a + h_1, b + h_2) - (f(a, b) + T(h_1, h_2))|}{|(h_1, h_2)|} = 0.$$

Rewriting this in terms of g, we obtain

$$\lim_{h_1, h_2 \to 0} \frac{|g(a + h_1) - (g(a) + T(h_1, h_2))|}{|(h_1, h_2)|} = 0,$$

which is satisfied by $T(h_1, h_2) = g'(a) \cdot h_1$.

2-3. Define when a function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is independent of the first variable and find $f'(a,b)$ for such f. Which functions are independent of the first variable and also of the second variable?

This is the same as the previous problem. A function f is independent of the first variable iff $f(x,y) = g(y)$ for some $g : \mathbb{R} \longrightarrow \mathbb{R}$. The derivative $f'(a,b)$ is the linear map $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ given by

$$T(h_1, h_2) = g'(b) \cdot h_2.$$

The functions $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ that are independent of both the first and second variables are the constant functions.

2-4. Let g be a continuous real-valued function on the unit circle $\{x \in \mathbb{R}^2 : |x| = 1\}$ such that $g(0,1) = g(1,0) = 0$ and $g(-x) = -g(x)$. Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} |x| \cdot g(\frac{x}{|x|}) & x \neq 0, \\ 0 & x = 0. \end{cases}$$

(a) If $x \in \mathbb{R}^2$ and $h : \mathbb{R} \longrightarrow \mathbb{R}$ is defined by $h(t) = f(tx)$, show that h is differentiable.

We start by establishing that a subset of the conditions on f are equivalent to homogeneity, i.e., the property that $f(tx) = t \cdot f(x)$. First, note that the given f is homogenous. This follows from

$$\begin{aligned} |tx| g(\frac{tx}{|tx|}) &= |t| \cdot |x| g(\frac{t}{|t|} \cdot \frac{x}{|x|}) \\ &= |t| \cdot |x| g(\text{sign}(t) \cdot \frac{x}{|x|}) \\ &= |t| \text{sign}(t) \cdot |x| g(\frac{x}{|x|}), \text{ because } g(-u) = -g(u) \\ &= t \cdot |x| g(\frac{x}{|x|}) \end{aligned}$$

In the other direction, given a homogenous $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, note that for any non-zero $x \in \mathbb{R}^2$,

$$f(x) = f(|x| \frac{x}{|x|}) = |x| f(\frac{x}{|x|}) = |x| g(\frac{x}{|x|})$$

where g is the restriction of f to the unit circle. Also, because f is homogenous, $f(0) = 0$, and $g(-x) = -g(x)$. Thus the homogeneity of f is *equivalent* to the properties listed above, except for $g(0,1) = g(1,0) = 0$.

Because f is homogenous, $h(t) = t\alpha$ for some $\alpha \in \mathbb{R}$, and so h is differentiable. In particular, $h' = \alpha$.

(b) Show that f is not differentiable at $(0,0)$ unless $g = 0$. Hint: First show that $Df(0,0)$ would have to be 0 by considering (h,k) with $k = 0$ and then with $h = 0$.

If it exists, $Df(0,0)$ is a linear map $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ which satisfies

$$\lim_{u,v \to 0} \frac{|f(u,v) - T(u,v)|}{|(u,v)|} = 0.$$

Because it is a linear map, $T(u,v) = \alpha u + \beta v$ for $\alpha, \beta \in \mathbb{R}$. Following Spivak's hint, if $u = 0$, then $f(u,v) = 0$ for all v, and the argument to the limit becomes $|\beta v|/|v| = |\beta|$. For this to go to zero with v, we need $\beta = 0$. Likewise, if $v = 0$ then we need $\alpha = 0$.

Thus after considering only two ways of varying u and v, we have the requirement that $T = 0$. Clearly, if f is zero everywhere (which happens iff g is), then the above limit is satisfied at the origin, making f differentiable there with derivative $T = 0$. Consider however the case in which $g = g_1 \neq 0$ at some point x on the unit circle. Let $u = \epsilon x_1$ and $v = \epsilon x_2$, so that as ϵ goes from 1 to 0, (u,v) goes from x to the origin. Note that $|f(u,v)| = |\epsilon g_1|$, and so the argument to the limit becomes

$$\frac{|f(u,v) - T(u,v)|}{|(u,v)|} = \frac{|\epsilon g_1|}{|\epsilon|} = |g_1|,$$

which does not go to zero with ϵ. Therefore, in this case, f is not differentiable at the origin.

2-5. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}} & (x,y) \neq 0, \\ 0 & (x,y) = 0. \end{cases}$$

Show that f is a function of the kind considered in Problem 2-4, so that f is not differentiable at $(0,0)$.

In part (a) of Problem 2-4, we showed that the main hypotheses on f are equivalent to homogeneity, i.e., that $f(\alpha x) = \alpha \cdot f(x)$.

Homogeneity follows easily for f in the current problem. The additional hypothesis from Problem 2-4 is also immediate, namely that $g(0,1) = g(1,0) = 0$. Then, because $g \neq 0$, we know from Problem 2-4 (b) that f is not differentiable at $(0,0)$.

2-6. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined by $f(x, y) = \sqrt{|xy|}$. Show that f is not differentiable at $(0, 0)$.

This is a repeat of part (b) from Problem 2-4. If the derivative of f exists at $(0, 0)$, it is the linear map $T \in (\mathbb{R}^2, \mathbb{R})$ which satisfies

$$\lim_{x,y \to 0} \frac{|f(x, y) - T(x, y)|}{|(x, y)|} = 0.$$

By setting $x = 0$ and varying y, and then setting $y = 0$ and varying x, it follows that $T = 0$. However if $x = y = \epsilon$, then $f(x, y) = |\epsilon|$, and so

$$\frac{|f(x, y) - T(x, y)|}{|(x, y)|} = \frac{|\epsilon|}{|\epsilon|\sqrt{2}} = \frac{1}{\sqrt{2}},$$

which does not approach 0 with x and y. It follows that f is not differentiable at $(0, 0)$.

2-7. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function such that $|f(x)| \leq |x|^2$. Show that f is differentiable at 0.

For f to be differentiable at 0 we need a linear map $T \in (\mathbb{R}^n, \mathbb{R})$ for which

$$\lim_{x \to 0} \frac{|f(x) - (f(0) + T(x))|}{|x|} = 0.$$

Let's try $T = 0$. Note that $|f(x)| \leq |x|^2$ implies $f(0) = 0$, and so we require

$$\lim_{x \to 0} \frac{|f(x)|}{|x|} = 0.$$

But this follows from the hypothesis on f. In particular, because

$$\frac{|f(x)|}{|x|} \leq \frac{|x|^2}{|x|} = |x|.$$

2-8. Let $f : \mathbb{R} \longrightarrow \mathbb{R}^2$. Prove that f is differentiable at $a \in \mathbb{R}$ if and only if f^1 and f^2 are, and that in this case

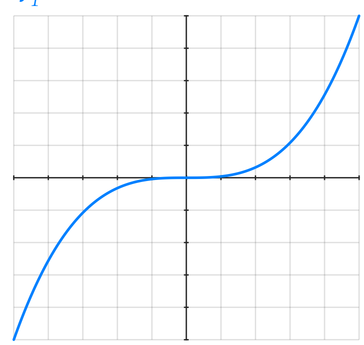
$$f'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \end{pmatrix}.$$

For f to be differentiable at a, a linear map T needs to exist which causes the following quotient to go to zero with h.

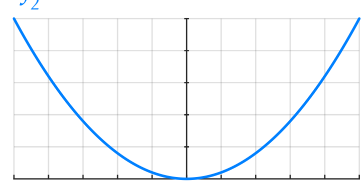
$$Q = \frac{|Th - (f(a+h) - f(a))|}{|h|}$$

Take a moment to consider what f might stand for. It could represent the time evolving frequency and amplitude of a radio wave, or the shape of a curve embedded in the plane, or many other things. In the case of a planar curve, f_1 and f_2 (and therefore f as we show in this exercise) can be smooth and yet leave the curve with a cusp!

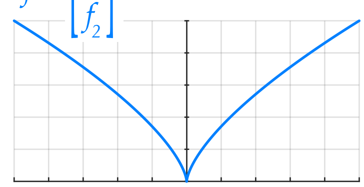
$f_1 = x^3$



$f_2 = x^2$



$f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$



The map T takes $h \in \mathbb{R}$ and returns $[\alpha h \;\; \beta h]^T \in \mathbb{R}^2$ for some $\alpha, \beta \in \mathbb{R}$. Expanding the terms in Q, we obtain

$$Q = \left| \begin{bmatrix} \alpha h \\ \beta h \end{bmatrix} - \left(\begin{bmatrix} f_1(a+h) \\ f_2(a+h) \end{bmatrix} - \begin{bmatrix} f_1(a) \\ f_2(a) \end{bmatrix} \right) \right| \frac{1}{|h|} = \left| \begin{bmatrix} A \\ B \end{bmatrix} \right| \frac{1}{|h|},$$

where $A, B \in \mathbb{R}$ denote the components in the numerator.

Suppose the components of f are differentiable at a. We prescribe T in terms of these derivatives. Setting $\alpha = Df_1(a)$ causes

$$\frac{|A|}{|h|} = \frac{|\alpha h - (f_1(a+h) - f_1(a))|}{|h|} \to 0 \text{ as } h \to 0.$$

Likewise setting $\beta = Df_2(a)$ causes $|B|/|h|$ to go to zero with h. In Problem 1-1 we showed that $|x| \leq \sum |x_i|$, and so

$$Q = \left| \begin{bmatrix} A \\ B \end{bmatrix} \right| \frac{1}{|h|} \leq \frac{|A|}{|h|} + \frac{|B|}{|h|}.$$

Our selection of α and β therefore causes $Q \to 0$ as $h \to 0$. It follows that f is differentiable at a as desired.

Conversely, if f is differentiable at a, then because $|x_i| \leq |x|$, we have

$$\frac{|A|}{|h|} \leq \left| \begin{bmatrix} A \\ B \end{bmatrix} \right| \frac{1}{|h|} = Q$$

Thus $|A|/|h| \to 0$ as $h \to 0$. It follows that $f_1(x)$ is differentiable at a, with derivative given by the α from T. Similarly, $Df_2(a) = \beta$.

2-9. Two functions $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ are **equal up to nth order** at a if

$$\lim_{h \to 0} \frac{f(a+h) - g(a+h)}{h^n} = 0.$$

(a) Show that f is differentiable at a if and only if there is a function g of the form $g(x) = a_0 + a_1(x-a)$ such that f and g are equal up to first order at a.

If f is differentiable at a, then some $\beta \in \mathbb{R}$ exists such that

$$\lim_{h \to 0} \frac{|f(a+h) - (f(a) + \beta h)|}{|h|} = 0.$$

If $g(x) = f(a) + \beta(x-a)$, then $g(a+h) = f(a) + \beta h$, and

$$\lim_{h \to 0} \frac{|f(a+h) - g(a+h)|}{|h|} = 0.$$

Finally, $A \to 0$ iff $|A| \to 0$, and so

$$\lim_{h \to 0} \frac{f(a+h) - g(a+h)}{h} = 0, \tag{3}$$

which means that f and g are equal up to first order at a.

The converse as stated is false. This is because a limit concerns the behavior *on approach* to a point, rather than at a point. The limit of $f(x)$ as x approaches a is independent of the value of $f(a)$, or whether f is even defined at a. Spivak is clear about this in the last section of Chapter 1. As a result, equation (3) can be true with f taking on *any* value at a, and so f doesn't even need to be continuous at a.

The converse is true if we add the condition that $f(a) = g(a)$. If $g = a_0 + a_1(x-a)$, then this condition causes $a_0 = f(a)$. Then (3) and the fact that $A \to 0$ iff $|A| \to 0$ gives us

A different way to make the converse true is to add the condition that f be continuous at a.

$$\lim_{h \to 0} \frac{|f(a+h) - (f(a) + a_1 h)|}{|h|} = 0.$$

This is the definition of f being differentiable at a, with derivative $Df(a) = a_1$.

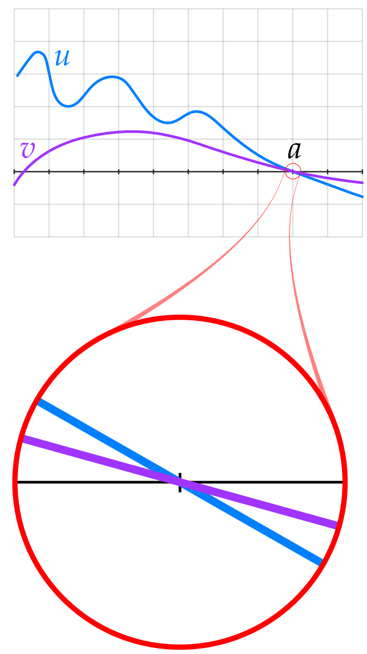
(b) If $f'(a), ..., f^{(n)}(a)$ exist, show that f and the function g defined by

$$g(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i$$

are equal up to nth order at a. Hint: The limit

$$\lim_{x\to a} \frac{f(x) - \sum_{i=0}^{n-1} \frac{f^{(i)}(a)}{i!}(x-a)^i}{(x-a)^n}$$

may be evaluated by L'Hospital's rule.



One of my fond memories from university is of being admonished during an early morning lecture to treat L'Hospital's rule with respect. "It is like a race car!" my professor exclaimed, grasping for a way to connect with sleepy 18 year olds. "It is capable of impressive speed, but also of impressive crashes."

I still think of a race car every time L'Hospital's rule comes up, however I also think of the picture to the right. If $u(x)$ and $v(x)$ both approach 0 as x approaches a, and if both are differentiable there, then u and v both look linear as you zoom in on a neighborhood of a. If $u(x) = \alpha x$ and $v(x) = \beta x$, then by inspection (or basic trig if you like), we obtain L'Hospital's rule:

$$\lim_{x\to 0} \frac{u(x)}{v(x)} = \lim_{x\to 0} \frac{u'(x)}{v'(x)} = \frac{\alpha}{\beta}.$$

In the given problem, for f and g to be equal up to nth order at a, we need

$$\lim_{h\to 0} \frac{f(a+h) - g(a+h)}{h^n} = 0.$$

Expanding the numerator $U(h)$, we obtain

$$U(h) = f(a+h) - (A_0 + A_1 h + A_2 \frac{h^2}{2!} + A_3 \frac{h^3}{3!} + ... + A_n \frac{h^n}{n!}),$$

where $A_k = f^{[k]}(a)$. Note that $U^{[k]}(0) = 0$ for $k = 0, 1, 2, ..., n$.

Next, consider the denominator $V(h) = h^n$. Note that $V^{[k]}(0) = 0$ for $k = 0, 1, 2, ..., n-1$, but that $V^{[n]}(0) = n!$. Thus the desired limit can be evaluated by $n-1$ applications of L'Hospital's rule, the last of which gives a quotient that is equal to zero. It follows that f and g are equal up to nth order at a.