# Math 338: Homework 7

Please complete the exercises below and write up your solutions consistent with the directions in the syllabus. Your solutions are due on Saturday, April 20th in the appropriate box outside my office door. If you get stuck on any part of the homework, please come and see me. More importantly, have fun!

## 1 Single-variable differentiation

**Exercise 1.** Below is a list of functions mapping real numbers to real numbers (though sometimes the domain will be a proper subset of real numbers). For each function:

a. Determine the function's natural domain.

b. Identify the difference quotient and its domain. Note: Rudin uses $\phi$ and I use $\Delta_f(t;x)$. Whichever one you decide to use is fine with me.

c. Using the definition in terms of the difference quotient, determine where (within its domain) the function is differentiable, in this case, determine the value of the derivative.

d. Using the definition in terms of the difference quotient, determine where the function is not differentiable within its domain.

Of course, your arguments should be rigorous. You can use anything in Chapters 1-4 of Rudin, but please don't use anything beyond the very first definition in Chapter 5 of Rudin.

1. Given $m, b \in \mathbb{R}$, consider the "affine" function $T(x) = mx + b$ for $x \in \mathbb{R}$.
2. For $C \neq 0$ and $\alpha \in \mathbb{Z}$, $p(x) = Cx^\alpha$.
3. $x \mapsto \sqrt{x}$.
4. $x \mapsto |x|$.

**Exercise 2.** In class I stated the following characterization of differentiability:

**Theorem 1.** *Let $I$ be an interval in $\mathbb{R}$, $f : I \to \mathbb{R}$, and $x \in I$. Then $f$ is differentiable at $x$ if and only if there is a linear function $L : \mathbb{R} \to \mathbb{R}$ (which is necessarily of the form $L(h) = Dh$ where $D = (m)$ is a $1 \times 1$ matrix with its single entry $m \in \mathbb{R}$) for which*

$$f(x+h) - f(x) - L(h) = \mathcal{E}(h)\,|h| \tag{1}$$

*where $\mathcal{E}$ is a real-valued function having the property that*

$$\lim_{h \to 0} \mathcal{E}(h) = 0. \tag{2}$$

*In this case, $L(h) = Dh$ where $D = (f'(x))$.*

Prove the theorem. *Hint: If you first assume that $f$ is differentiable at $x$ (using the definition at the beginning of Chapter 5 of Rudin), then define $L(h) = (f'(x))h = f'(x)h$ and*

$$\mathcal{E}(h) = \frac{f(x+h) - f(x) - L(h)}{|h|}$$

*for $h$ such that $x + h \in I$. Your job is then to show that $\mathcal{E}$ satisfies (1) and (2). Conversely, assume that you have a linear map $L(h) = (m)h$ and $\mathcal{E}$ satisfying (1) and (2). Then, you must show that $f$ is differentiable and $m = f'(x)$. In the course of your argument, you might find the need the following lemma:*

**Lemma 2.** *Let $g$ be a real valued function defined on an interval containing $0$. Then*

$$\lim_{h\to 0} g(h) = 0 \qquad \text{if and only if} \qquad \lim_{h\to 0} |g(h)| = 0.$$

*If you use the lemma, you must also prove it.*

**Exercise 3.** The chain rule for single-variable function is the following.

**Theorem 3.** *Given intervals $I$ and $J$, let $f : I \to \mathbb{R}$ be such that $f(I) \subseteq J$ and let $g : J \to \mathbb{R}$. We define the composition $g \circ f : I \to \mathbb{R}$ by*

$$g \circ f(x) = g(f(x))$$

*for $x \in I$. If $f$ is differentiable at $x$ and $g$ is differentiable at $y = f(x)$, then $g \circ f$ is differentiable at $x$ and*

$$(g \circ f)'(x) = g'(f(x))f'(x).$$

1. Use Theorem 1 (directly) to prove the chain rule. *Note/Hint: Your proof should be easier than (and distinct from) Rudin's. If it isn't, something is going wrong and you should come and chat with me. To get started, write out* (1) *and* (2) *for $f$ at $x$ and $g$ at $y = f(x)$ (making sure to label the $\mathcal{E}'$s differently). Then put them together.*

2. Use the chain rule (and the definition of the derivative) to differentiate the following functions (including saying where they are differentiable and finding their derivative everywhere it is possible).

   (a)
   $$h(x) = \begin{cases} x^2\cos(1/x) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

   (b)
   $$j(x) = e^{e^{e^{e^{e^{x}}}}}$$

   *Hint: Though we have not proved it, you may assume that the exponential and cosine functions have the derivative you know them to have.*

## 2 Multivariate Calculus

In what follows, I will discuss the continuity and differentiability of functions mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. To make things as clear as possible, points/vectors in $\mathbb{R}^n$ (or $\mathbb{R}^m$) will be expressed henceforth as column vectors and, when not written out in components/coordinates, I will write them in boldface. For instance,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

and, further, the (Euclidean) norm of this vector is

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

We should remark that, in $\mathbb{R}^n$, the Euclidean metric $d = d_2 = d_{\mathbb{R}^n}$ is given by $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. On Friday, we investigated differentiability for functions mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Recall that, every linear function $L : \mathbb{R}^n \to \mathbb{R}^m$ can be written (uniquely) in the form

$$L(\mathbf{x}) = D\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{for} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

where $D$ is the $m \times n$ matrix having entries $a_{ij}$ as indicated above. Here is the definition of differentiability:

**Definition 4.** Given an open set $\mathcal{O} \subseteq \mathbb{R}^n$, let $f : \mathcal{O} \to \mathbb{R}^m$; here $n$ and $m$ are positive integers. Given a point $\mathbf{x}_0 \in \mathcal{O}$, we say that $f$ is differentiable at $\mathbf{x}_0$ if there is a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ for which

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - L(\mathbf{h}) = \mathcal{E}(\mathbf{h})\|\mathbf{h}\| \tag{3}$$

where $\mathcal{E}$ (which depends on $\mathbf{x_0}$ and $f$) is an $\mathbb{R}^m$-valued function having the property that

$$\lim_{\mathbf{h}\to\mathbf{0}} \mathcal{E}(\mathbf{h}) = \mathbf{0}. \tag{4}$$

In this case, the **derivative** of $f$ at $\mathbf{x}_0$ (or **Jacobian derivative** or **Total Derivative**) is the $m \times n$ matrix $D = Df(\mathbf{x}_0)$ for which $L(\mathbf{h}) := D\mathbf{h}$ for $\mathbf{h} \in \mathbb{R}^n$.

Before moving on to more exercises, let's make two remarks.

*Remark* 1. If we translate what the limit means in terms of Euclidean metrics on $\mathbb{R}^n$ and $\mathbb{R}^m$, the limit (4) is equivalently the statement: For every $\epsilon > 0$, there is a $\delta > 0$ for which

$$\|\mathcal{E}(\mathbf{h})\| < \epsilon \qquad \text{whenever} \qquad \|\mathbf{h}\| < \delta;$$

here, you have to a little careful because the first appearance of $\|\cdot\|$ means the norm on $\mathbb{R}^m$ and the second is the norm on $\mathbb{R}^n$. If it's helpful to you, we established an equivalence between the 1, 2, and $\infty$ metrics/norms on Homework 2 that you can freely use.

*Remark* 2. In the definition above, I use the definite article in the phrase ".. .the $m \times n$ matrix..." As we discussed in Homework 1, when we do this, we should make sure that such a matrix is indeed unique. This is the subject of Exercise 5 in the present homework. Outside of that exercise, you can take it for granted.

**Exercise 4.** Let's verify differentiability for some functions using the definition above. In what follows, I give three functions. For each function, I will identify a point in the functions domain and a candidate for its derivative at that point. Please use the definition to show that each function is differentiable at the indicated point.

1. $f : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$f(\mathbf{x}) = \begin{pmatrix} xyz \\ x+z \end{pmatrix} \qquad \text{for} \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

Here

$$\mathbf{x}_0 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

2. $g : \mathbb{R} \to \mathbb{R}^2$ defined by

$$g(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

for $t \in \mathbb{R}$. The point at which I'd like you to investigate is an arbitrary point $t_0 \in \mathbb{R}$ and, for this point, the candidate for the derivative is

$$D = \begin{pmatrix} -\sin(t_0) \\ \cos(t_0) \end{pmatrix}.$$

Here, you may freely use any trigonometric identities and the following two facts: $|\cos(u) - 1| \leq u^2$ and $|\sin(u) - u| \leq u^3$ whenever $|u| \leq 1$. We will prove these at some point.

3. Let $h : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$h\begin{pmatrix} x \\ y \end{pmatrix} = xy \qquad \text{for} \qquad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Here, for some $x_0, y_0 \in \mathbb{R}$,

$$\mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} y_0 & x_0 \end{pmatrix}.$$

**Exercise 5.** In this exercise, we shall prove that our derivatives (and affine approximations) to differentiable functions are unique. An affine function $T$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is, by definition, a function of the form

$$T(\mathbf{x}) = L(\mathbf{x}) + \mathbf{b}$$

where $L(\mathbf{x}) = D\mathbf{x}$ is a linear function (characterized by the $m \times n$ matrix $D$) and $\mathbf{b} \in \mathbb{R}^m$. Using this language of affine functions, we can state differentiability of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ as follows: $f$ is differentiable at a point $\mathbf{x}_0$ if there is an affine function $T : \mathbb{R}^n \to \mathbb{R}^m$ for which

$$f(\mathbf{x}_0 + \mathbf{h}) = T(\mathbf{h}) + \mathcal{E}(\mathbf{h})\|\mathbf{h}\| \tag{5}$$

for $\mathbf{h} \in \mathbb{R}^n$ where $\lim_{h\to\mathbf{0}} \mathcal{E}(\mathbf{h}) = \mathbf{0}$. Prove that, if $f$ is differentiable at $\mathbf{x}_0$, then there is one and only one affine function $T$ for which (5) holds. *Hint: Differentiability should give you one affine function. After that, your job is to assume that, if you have any two such functions, they must be equal. The following lemma might be helpful. If you use it, give a proof.*

**Lemma 5.** *Let $M : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Then $M$ is not the zero transformation if and only if there is some vector $\mathbf{v} \in \mathbb{R}^n$ with unit length (i.e., $\|\mathbf{v}\| = 1$) for which $M(\mathbf{v}) \neq \mathbf{0}$.*

**Exercise 6.** Let $\mathcal{O}$ be open in $\mathbb{R}^n$ and let $f : \mathcal{O} \to \mathbb{R}$. Given a point $\mathbf{x}_0 \in \mathcal{O}$ and an index $j \in \{1, 2, \dots, n\}$, the partial derivative of $f$ with respect to the variable $x_j$ at $\mathbf{x}_0$ is defined by

$$\partial_j f(\mathbf{x}_0) = \frac{\partial f}{\partial x_j}(\mathbf{x}_0) = \lim_{h\to 0} \frac{f(\mathbf{x}_0 + h\mathbf{e}_j) - f(\mathbf{x}_0)}{h}$$

provided this limit exists; here, $\mathbf{e}_j \in \mathbb{R}^n$ is the unit vector with 1 in the $j$th entry and zeros everywhere else, i.e.,

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \ \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

1. Directly using the definition, compute all (first-order) partial derivatives of the functions (here, I'm using the convention that $x_1 = x$, $x_2 = y$ and $x_3 = z$):
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto xyz \qquad \text{and} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x + z.$$
2. Let $\mathcal{O} \subseteq \mathbb{R}^n$ and $f : \mathcal{O} \to \mathbb{R}$. Given a point $\mathbf{x}_0 \in \mathcal{O}$ and an index $j \in \{1, 2, \dots, n\}$, define
$$g_j(t) = f(\mathbf{x}_0 + t\mathbf{e}_j).$$
Show that $g_j$ is differentiable at 0 if and only if the partial derivative of $f$ with respect to $x_j$ at $\mathbf{x}_0$ exists and, in this case,
$$g_j'(0) = \frac{\partial f}{\partial x_j}(\mathbf{x}_0).$$
3. It you look closely you'll see that $g_j$ above is the composition of functions, the inner such function being vector-valued and differentiable everywhere. Use the chain rule to prove the following theorem:

   **Theorem 6.** *Let $f : \mathcal{O} \to \mathbb{R}$ be as above. If $f$ is differentiable at $\mathbf{x}_0$, then all of its (first-order) partial derivatives $\partial f/\partial x_1$, $\partial f/\partial x_2$, ... $\partial f/\partial x_n$ exists at $\mathbf{x}_0$ and, moreover,*
$$Df(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$$
   We'll have a more general version of this in class on Monday.
4. Go back and confirm that the above formula (which determines the derivative matrix) is consistent with the proposed derivative matrices in Exercise 4.