MATH20410 (W25): Analysis in Rn II (accelerated)

Lecturer: Joe Jackson
Notes by: Aden Chen

Wednesday 22nd January, 2025

Contents

1 Single-Variable Differential Calculus

In this chapter, we consider mainly functions of the form $f : I \to \mathbb{R}$, where I is an interval, e.g., (a, b), $[a, b]$, $(a, b]$, (a, ∞), $\mathbb{R}$. This is the function we have in mind unless otherwise stated.

Definition 1.1 (Differentiability)**.** We say f is **differentiable at** $x \in I$ if the limit

$$f'(x) \coloneqq \lim_{t \to x} \frac{f(t) - f(x)}{t - x} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

exists. In this case, we call $f'(x)$ the derivative of f at x. Moreover:

- We say that f is **differentiable** if $f'(x)$ exists for each $x \in I$.
- We say f is **continuously differentiable** ($f \in C^1$) if $f' : I \to \mathbb{R}$ is continuous.

Example 1.2.

- $f(x) = |x|$. Differentiable on $\mathbb{R} \setminus \{0\}$.
- $f(x) = \begin{cases} x \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$. Continuous but not differentiable at 0.
- $f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$. Differentiable everywhere (in particular at 0), but $f \notin C^1$.

Proposition 1.3 (Rules for computing derivatives)**.**

(i) Linearity. $(af + bg)' = af' + bg'$ *(if f' and g' exist, such requirements are hereafter omitted).*

(ii) Product rule. $(fg)' = f'g + fg'$.

(iii) Quotient rule. $(f/g)' = (f'g - fg')/g^2$.[1]

(iv) Chain rule. $(f \circ g)' = (f' \circ g) \cdot g'$.

[1] Low dhigh minus high dlow. Not Haidilao...

Proof. We prove the quotient rule; the remaining are left as exercises. Starting from the definition

$$\left(\frac{f}{g}\right)'(x) = \lim_{t\to x} \frac{\frac{f}{g}(t) - \frac{f}{g}(x)}{t-x}$$
$$= \lim_{t\to x} \frac{\frac{f(t)}{f(t)} + \frac{f(x)}{g(t)} - \frac{f(x)}{g(t)} + \frac{f(x)}{g(x)}}{t-x}.$$

Note that

$$\frac{\frac{f(x)}{g(t)} + \frac{f(x)}{g(x)}}{t-x} = \frac{f(x)}{g(x)g(t)} \frac{g(x)-g(t)}{t-x}$$

and we have

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{g^2(x)}$$

□

Theorem 1.4. *If f is differentiable at x then f is continuous at x.*

Proof. Note that

$$\lim_{t\to x} f(t) - f(x) = \lim_{t\to x} \frac{f(t)-f(x)}{t-x}(t-x) = f'(x)\cdot 0 = 0.$$

□

1.1 The Mean Value Theorem

Lemma 1.5. *Suppose $f : [a,b] \to \mathbb{R}$ has a local maximum or minimum at $x \in (a,b)$. If $f'(x)$ exists, then $f'(x) = 0$.*

Proof. From the definition of the derivative, consider the limits from the left and right; one is non-positive and the other is non-negative. □

Theorem 1.6 (Rolle's Theorem)**.** *Suppose $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$, differentiable on (a,b), and such that $f(a) = f(b)$. Then there exists $x \in (a,b)$ such that $f'(x) = 0$.*

Proof. Consider the global maximum or minimum (exist since f is a continuous function defined on a compact set) and apply the previous lemma. (If both the maximum and minimum is at a or b, f is constant.) □

Theorem 1.7 (Mean Value Theorem)**.** *Let $f : [a,b] \to \mathbb{R}$ be such that f is continuous on $[a,b]$ and differentiable on (a,b). Then there exists $x \in (a,b)$ such that $f(b) - f(a) = f'(x)(b-a)$.*

Proof. Apply Rolle's to $\tilde{f} = f - [f(b) - f(a)] \cdot \frac{x-a}{b-a}$. □

Theorem 1.8. *Let $f : (a,b) \to \mathbb{R}$ be differentiable.*

(a) if $f' = 0$, then f is constant.

(b) if $f' \geq 0$, then f is increasing.

(c) if $f' \leq 0$, then f is decreasing.

Proof. Apply the mean value theorem. □

Theorem 1.9 (The Intermediate Value Property of Derivatives)**.** *Let $f : [a,b] \to \mathbb{R}$ be differentiable*[2] *and suppose $f'(a) < \lambda < f'(b)$ Then there exists $x \in (a,b)$ such that $f'(x) = \lambda$.*

[2] f need not be C^1!

Proof *(à la Pugh)***.** Slide a small secant of length so small that the slope around a and b is separated also by λ. By continuity of the slope, there exists a secant between a and b with slope λ. Apply the mean value theorem to this slope. □

Proof *(à la Joe/Rudin)***.** We start with $\lambda = 0$. Then $f'(a), f'(b) \neq 0$ and the global min/max of f cannot be at the endpoints. At the global extrema we have the desired result. When $\lambda \neq 0$, consider $\tilde{f} \coloneqq f - \lambda x$. □

Example 1.10. Consider

$$f(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

We have

$$f(x) = \begin{cases} 2x \sin(1/x) = \cos(1/x), & x \neq 0 \\ 0, & x = 0 \end{cases},$$

which has the intermediate value property.

Theorem 1.11 (Generalized Mean Value Theorem)**.** *Let $f, g : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then there exists $x \in (a,b)$ such that*

$$\big(f(a) - f(b)\big)g'(x) = \big(g(a) - g(b)\big)f'(x).$$

Remark 1.12. When the above is not zero,

$$\frac{f(a) - f(b)}{g(a) - g(b)} = \frac{f'(x)}{g'(x)}.$$

Proof. Define

$$h(t) := \big(f(b) - f(a)\big)g(t) - \big(g(b) - g(a)\big)f(t).$$

Note that

$$h(a) = f(b)g(a) - f(a)g(b) = h(b)$$

and apply Rolle's. $\square$

1.2 L'Hôpital's Rule

Theorem 1.13 (L'Hôpital's Rule, a particular case)**.** *Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $g(x) \neq 0$ in a neighborhood of a and $f(x) = g(x) = 0$, then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

if the last limit exists.

Proof. Consider some small $\delta > 0$. The generalized MVT gives some $x \in (a, a+\delta)$ such that

$$\frac{f(a+\delta)}{g(a+\delta)} = \frac{f'(x)}{g'(x)} \approx \lim_{t \to a} \frac{f'(t)}{g'(t)},$$

where the last approximation follows from the existence of the limit. Note that as $\delta \to 0$, $x \to a$, and the approximation error shrinks to 0. $\square$

Refer to Rudin or something for the general case.

1.3 Higher Derivatives

If $f : I \to \mathbb{R}$ is differentiable, then we can define the second derivative $f'' := (f')'$ if f' is differentiable. Higher derivatives can be defined similarly. We usually write $f^{(n)}$ for the n-th derivative of f.

Example 1.14. $L(x) = f(x_0) + f'(x_0)(x - x_0)$ is a (first order) linear approximation of f at x_0. How good is this approximation? A first answer is

$$f(x) = L(x) + o(|x - x_0|),$$

since we have as $x \longrightarrow x_0$ that

$$\frac{f(x) - L(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \longrightarrow 0.$$

But can we say even more about the quality of the approximation? – Yes, if f is twice differentiable.

Proposition 1.15 (First-order Taylor's Theorem)**.** *Suppose f' exists and is continuous on $[a, b]$ and f'' exists on (a, b). Let $x_0, x \in [a, b]$ with $x_0 \neq x$. Then*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(y)(x - x_0)^2,$$

where y is between x_0 and x. In particular, we have

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq \frac{1}{2} \sup_{y \in (a,b)} |f''(y)| \cdot |x - x_0|^2.$$

Proof. Find M such that we have

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{M}{2}(x - x_0)^2.$$

We need only find y such that $M = f''(y)$. Define

$$g(t) \coloneqq f(t) - f(x_0) - f'(x_0)(t - x_0) - \frac{M}{2}(t - x_0)^2.$$

Note that $g''(t) = f''(t) - M$, so we need only find a point at which g'' vanishes. Since $g(x_0) = g(x) = 0$, by the MVT there exists y' between x_0 and x such that $g(y') = 0$. Observe that $g'(x_0) = 0$, and so by the MVT again, there exists y between x_0 and y' (and by extension between x_0 and x) such that $g''(y) = 0$. $\square$

The more general story: given $f : [a, b] \longrightarrow \mathbb{R}$ and $x_0 \in [a, b]$, we may define

$$\begin{aligned}
P_0(x) &\coloneqq f(x_0), \\
P_1(x) &\coloneqq f(x_0) + f'(x_0)(x - x_0), \\
P_2(x) &\coloneqq f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2, \\
&\vdots \\
P_n(x) &\coloneqq \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k,
\end{aligned}$$

when the corresponding derivatives exist. Note that $P_n(x)$ is the unique degree n polynomial such that $P_n^{(k)}(x_0) = f^{(k)}(x_0)$ for $k = 1, \ldots, n$.

Theorem 1.16 (Taylor's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be such that*

- *$f^{(k)}$ exists on $[a, b]$ for $k = 1, \ldots, n$; and*
- *$f^{(n+1)}$ exists on (a, b).*

Then, for any $x_0, x \in [a, b]$ with $x_0 \neq x$, there exists y between x_0 and x such that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(y)}{(n+1)!}(x - x_0)^{n+1}.$$

for some y between x_0 and x.

We proof the case $n = 2$, the same idea can be used to prove the general case.

Proof. Define

$$g(t) = f(t) - P_2(t) - \frac{M}{6}(t - x_0)^3.$$

Since $g''' = f''' - M$, we need only find y such that $g'''(y) = 0$. Note that $g(x_0) = g(x) = 0$, and so by the MVT there exists y' between x_0 and x such that $g'(y') = 0$. Next, note that $g'(x_0) = 0$, and so by the MVT there exists y'' between x_0 and y' such that $g''(y'') = 0$. Finally, note that $g''(x_0) = 0$, and so by the MVT there exists y between x_0 and y'' such that $g'''(y) = 0$. $\square$

2 Multivariable Differential Calculus

Some remainders about $\mathbb{R}^n$:

- $\mathbb{R}^n = \{x = (x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$.
- $\mathbb{R}^n$ is a vector space, with canonical basis $\{e_i, \ldots, e_n\}$.
- $\mathbb{R}^n$ comes with an inner product $\langle x, y\rangle = x \cdot y = \sum x_i y_i$, a norm $|x| = \sqrt{x \cdot x} = (\sum x_i y_i)^{1/2}$, and a metric $d(x, y) = |x - y|$.

2.1 Higher Dimensional Codomains

Consider a function $f : \mathbb{R} \supset I \to \mathbb{R}^n$.

Definition 2.1. f is differentiable at x if the limit

$$f'(x) \coloneqq \lim_{t\to x} \frac{f(t) - f(x)}{t - x}$$

exists.

Remark 2.2. We may write $f(t) = (f_1(t), \ldots, f_n(t))$, and $f'(x) = (f_1'(x), \ldots, f_n'(x))$, since a sequence $x \in \mathbb{R}^n$ converges if and only if each of its components converges.

Theorem 2.3. *We have the following analog of the MVT:*

$$|f(b) - f(a)| \leq |f'(t)| \cdot |b - a|.$$

for some t between a and b.

Proof. Assume $a < b$. Define

$$h(t) \coloneqq \langle f(b) - f(a), f(t)\rangle .$$

The MVT gives

$$\begin{aligned} h(b) - h(a) = h'(t)(b - a) &= \langle f(b) - f(a), f'(t)\rangle (b - a) \\ &\leq (b - a)|f(b) - f(a)||f'(t)|, \end{aligned}$$

where the last inequality follows from the Cauchy-Schwarz inequality. Noting that

$$h(b) - h(a) = |f(b) - f(a)|^2,$$

we have the desired result. $\square$

2.2 Higher Dimensional Domain

We next consider functions $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ is open.

Definition 2.4 (Partial Derivatives)**.**

$$\frac{\partial f}{\partial x_i}(x) = \mathrm{D}_i f(x) \coloneqq \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h}.$$

Definition 2.5 (Directional Derivatives)**.** Fix $u \in \mathbb{R}^n$.

$$= \mathrm{D}_i u f(x) \coloneqq \lim_{h \to 0} \frac{f(x + hu) - f(x)}{h}.$$

2.2.1 The Derivative

Intuition: A function is differentiable if a first-order Taylor expansion holds. That is, if f is "well-approximated" by a linear function.

Definition 2.6. We denote the set of all linear maps from $\mathbb{R}^n$ to $\mathbb{R}$ as $L(\mathbb{R}^n, \mathbb{R})$.

Definition 2.7 (The Derivative)**.** A function f is differentiable at x if there exists a linear map $T \in L(\mathbb{R}^n, \mathbb{R})$ such that

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

In this case we write $\mathrm{D}f(x) = T$. In other words, $f(x + h) = f(x) + \mathrm{D}f(x)(h) + o(|h|)$.

Remark 2.8.

- If f is differentiable, then

$$\mathrm{D}f : U \longrightarrow L(\mathbb{R}^n, \mathbb{R}).$$

- If is easy to check that $\mathrm{D}f$ is well defined, that is, there is at most one T such that the limit holds.

We may think of the linear map $T : \mathbb{R}^n \to \mathbb{R}$ as

$$T(u) = \langle u, v \rangle, \tag{1}$$

where $v \coloneqq (Te_1, \ldots Te_n)$.

Definition 2.9 (The Gradient)**.** If f is differentiable at x, we define $\nabla f(x) = v$, where v satisfies (1). In other words,

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h \rangle}{|h|} = 0.$$

Theorem 2.10. *If f is differentiable at x, then $\mathrm{D}_u f(x)$ exists for all $u \in \mathbb{R}^n$ and $\mathrm{D}_u f(x) = \mathrm{D} f(x) u = \langle \nabla f(x), u \rangle$.*

Proof. Note that as $t \to 0$, we have

$$\begin{aligned} \left| \frac{f(x+tu) - f(x)}{t} - \mathrm{D} f(x) u \right| &= \left| \frac{f(x+tu) - f(x) - \mathrm{D} f(x)(tu)}{t} \right| \\ &= \left| \frac{f(x+tu) - f(x) - \mathrm{D} f(x)(tu)}{|tu|} \right| \cdot |u| \longrightarrow 0. \end{aligned}$$

□

Remark 2.11. In particular we have $\mathrm{D}_i f(x) = \mathrm{D}_{e_i} f(x) = \mathrm{D} f(x) e_i = \langle \nabla f(x), e_i \rangle$. In other words, if f is differentiable, then $\nabla f(x) = (\mathrm{D}_1 f, \ldots, \mathrm{D}_n f)$.

Remark 2.12.

- Differentiability holds if and only if the gradient exists.
- Differentiability implies the existence of directional derivatives, which then implies the existence of partial derivatives. The converse implications are not true.

Example 2.13. Consider

$$f(x_1, x_2) \coloneqq \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

It is easy to see that $\mathrm{D}_1 f(0) = \mathrm{D}_2 f(0) = 0$ but $\mathrm{D}_{(1,1)} f(0)$ does not exist. Indeed, f is not even continuous on the line $t(1,1)$.

Example 2.14. Consider

$$f(x_1, x_2) \coloneqq \begin{cases} \frac{x_1^3}{x_1^2 + x_2^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Note that

$$\mathrm{D}_u f(0) = \lim_{t \to 0} \frac{t^3 u_1^3}{t^2(u_1^2 + u_2^2)} \cdot \frac{1}{t} = \frac{u_1^3}{u_1^2 + u_2^2}.$$

However, $\mathrm{D}f(0)$ cannot exist, since the above mapping is not linear.

Theorem 2.15. *If the partial derivatives $\mathrm{D}_1 f, \ldots, \mathrm{D}_n f$ exist and are continuous (in a neighborhood of x), then f is differentiable at x.*

Proof. Fix arbitrary $x \in E$ and define $Ah = \sum \mathrm{D}_i f(x) h_i$. We write $\omega_k \coloneqq \sum_{i=1}^k h_i e_i$ for $k = 1, \ldots, n$ and $\omega_0 \coloneqq x$. Note that $\omega_n = h$. By the MVT we can find δ_k between 0 and h_k such that

$$\begin{aligned} f(x+h) - f(x) - Ah &= \sum_{k=1}^n f(x + \omega_k) - f(x + \omega_{k-1}) - \mathrm{D}_k f(x) h_k \\ &= \sum_{k=1}^n h_k [\mathrm{D}_k (x + \omega_k + \delta_i e_i) - \mathrm{D}_k f(x)], \end{aligned}$$

which by continuity of D_i is sublinear. □

2.3 Extension to Functions with Higher Dimensional Codomains

Immediate.

Note that we may identify $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ with a unique matrix $A = [Te_1, \ldots, Te_n]$ such that we have $Th = Ah$ for each h.