Some Analysis Things

Arthur Conmy*

Part IB, Easter Term 2021

In these notes, I make some brief comments on the IB Analysis courses[1]. Credit is due to Evan Chen for the style file for these notes[2].

§1 Integration

*Please send any corrections and/or feedback to asc70@cam.ac.uk.

[1] strictly, just Analysis and Topology and Complex Analysis, but I hoep that the reader agrees that Analysis and Topology is ... more than one courses worth of material!

[2] Available here: https://github.com/vEnhance/dotfiles/blob/master/texmf/tex/latex/evan/evan.sty.

Theorem 1.1 (Interchanging Differentiation and Integration)
Let $f : \mathbb{R} \times [0, 1] \to \mathbb{R}$ be a cts function of θ and t, and suppose $\frac{\partial f}{\partial \theta}$ is also cts.
Then

$$\frac{d}{d\theta} \int_0^1 f(\theta, t) dt = \int_0^1 \frac{\partial f}{\partial \theta}(\theta, t) dt. \tag{1}$$

Proof. The key idea is that differentiability is a local property, so we can force the domain of $\frac{\partial f}{\partial \theta}$ to be compact.

WLOG let $\theta = 0$. Now $\forall \varepsilon > 0$, pick $\delta > 0$ so that $\left|\frac{\partial f}{\partial \theta}(\theta, t) - \frac{\partial f}{\partial \theta}(0, t)\right| < \varepsilon$ for $(\theta, t) \in [-\delta, \delta] \times [0, 1]$. Then let

$$F(\theta) = \int_0^1 f(\theta, t) dt. \tag{2}$$

Then we can directly calculate (for $|h| < \delta$)

$$\frac{1}{h}(F(h) - F(0))) = \int_0^1 \frac{f(h, t) - f(0, t)}{h} dt = \int_0^1 \frac{\partial f}{\partial \theta}(\theta_t, t) dt. \tag{3}$$

where the last equality follows from the mean value theorem. $\theta_t \in (-|h|, |h|)$ is a function of t[a].

But our choice of delta means that this differs by at most ε from

$$\int_0^1 \frac{\partial f}{\partial \theta}(0, t) dt, \tag{4}$$

so we're done. □

[a]Extra: can we always make it a *cts* function of t?

Theorem 1.2 (CIF for Derivatives)
Let U be a domain and $f : U \to \mathbb{C}$ holomorphic. Let $D(0, 1) \subseteq U$ and $w \in D(0, 1)$. Then

$$f^{(n)}(w) = \frac{1}{n!} \oint_{\partial D(0,1)} \frac{f(z)}{(z - w)^{n+1}} dz. \tag{5}$$

Proof. Case $n = 1$ is the ordinary integral formula, and we show how the $n = 2$ arises by applying (**??**). The higher order cases arise similarly.

We can write the $n = 1$ case as

$$\oint_{\partial D(0,1)} \frac{f(z)}{z - w} dz = \int_0^1 \frac{f(\gamma(t))\gamma'(t)}{\gamma(t) - w} dt. \tag{6}$$

where $\gamma(t) = e^{2\pi i t}$. We can now directly apply the previous result, since the integrand's partial w derivative is indeed cts, provided we localise to a ball around w (so that the $\frac{1}{\gamma(t) - w}$ term doesn't get very large). □

§2 Differentiation

I don't think this part of the course needs to be anywhere near as feared as it is currently.

Definition 2.1 (Norm)**.** Let V be a real vector space. A norm on V is a function $||.|| : V \to \mathbb{R}$ satisfying

- $||v|| \geq 0$, with equality iff $v = 0$.
- $||v + w|| \leq ||v|| + ||w||$.
- $||\lambda v|| = |\lambda| ||v||$.

> **Remark 2.2.** This naturally induces a topology on V by turning it into a metric space with distance function $d(v, w) = ||v - w||$.

Theorem 2.3 (Only one norm)
Let V be a finite dimensional vector space. Then all norms on V are Lipschitz equivalent.

Proof. Fix a basis $e_1, ..., e_n$ of V. Then let $||.||_2$ be the Euclidean norm, i.e

$$||\lambda_1 e_1 + .. + \lambda_n e_n||_2 = \sqrt{\lambda_1^2 + ... + \lambda_n^2}. \tag{7}$$

We show that all norms are Lipschitz equivalent to the Euclidean norm, and since Lipschitz equivalence is an equivalence relation, this will suffice.

- $\exists m > 0$ such that $||v|| \leq m||v||_2$ for all $v \in V$:

  Direct application of the triangle inequality. Let

  $$E = \max_{i=1}^{n} ||e_i|| > 0. \tag{8}$$

  Then if $v = \lambda_1 e_1 + ... + \lambda_n e_n$, and

  $$\Lambda = \max_{i=1}^{n} |\lambda_i|, \tag{9}$$

  then

  $$||v|| \leq |\lambda_1|||e_1|| + ... + |\lambda_n|||e_n|| \leq E(|\lambda_1| + |\lambda_2| + ... + |\lambda_n|) \tag{10}$$

  where the first inequality follows from the triangle inequality. So

  $$||v|| \leq nE\Lambda. \tag{11}$$

  Now $||v||_2 \geq \Lambda$ and hence $m = nE$ works (note we need Λ independence, but E dependence is fine since the former is a property of the specific v, but the latter a property of the norm).

- $\exists M > 0$ such that $||v|| \geq M||v||_2$ for all $v \in V$:

  Less obvious.

$\square$

Why does this matter? Recall the definition of differentiablility

Definition 2.4. $f : \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at p with derivative the linear map $T(p)(h)$ if

$$\lim_{h:||h||_2 \to 0} \frac{||f(p+h) - f(p) - T(p)(h)||_2}{||h||_2}. \tag{12}$$

The important thing to notice here is that while f and $T(p)$ are both maps from $\mathbb{R}^m$, the former exclusively maps from points very close to p, and the latter exclusively maps from points very close to 0. If this isn't clear, chapter 42 of [?] (Napkin) provides a far better explanation than what I can give (and includes pictures!).

To be concrete however, working with a multi-dimensional limit is hard. In practice, we are likely to generally use the following result (cross-posted from my Geometry notes)

Theorem 2.5 (Computing derivatives in n dimensions.)
Suppose that $U \subseteq^\circ \mathbb{R}^m$ and $f : U \to \mathbb{R}$ has continuous partial derivatives at $p \in U$. Then f is differentiable at p, with derivative

$$Df(p)(h) = \frac{\partial f}{\partial x_1} h_1 + \cdots + \frac{\partial f}{\partial x_m} h_m. \tag{13}$$

Proof. Actually nowhere near as bad as some analyis proofs. Decompose as a telescoping sum

$$\begin{aligned}
& f(x_1 + h_1, \cdots, x_m + h_m) - f(x_1, \cdots, x_m) & (14)\\
= {} & f(x_1 + h_1, \cdots, x_m + h_m) - f(x_1 + h_1, \cdots, x_{m-1} + h_{m-1}, x_m) & (15)\\
+ {} & f(x_1 + h_1, \cdots, x_{m-1} + h_{m-1}, x_m) - f(x_1 + h_1, \cdots, x_{m-1}, x_m) & (16)\\
& + \cdots & (17)\\
& + f(x_1 + h_1, x_2, \cdots, x_m) - f(x_1, \cdots, x_m). & (18)
\end{aligned}$$

And now by an 'm-ε' proof (i.e m applications of the triangle inequality), we use continuity of the m partial derivatives to deduce the desired derivative expression. □

> **Remark 2.6.** This is basically the same sort of thing we did in IA DEs when we integrated from (x_1, y_1) to (x_2, y_2) by first travelling from (x_1, y_1) to (x_2, y_1) (with fixed y value) and then from (x_2, y_1) to (x_2, y_2) (with fixed x value), like a staircase.

> **Remark 2.7.** Actually, the last term in the telescoping series, $f(x_1 + h_1, x_2, \cdots, x_m) - f(x_1, \cdots, x_m)$ *is* the partial derivative (in x_1) of f at $(x_1, ..., x_n)$. So we only need continuity of $n - 1$ of the partial derivatives at a point, and existence of the last, in order to deduce differentiablility at a point.

Look again at (**??**). There's nothing special about $||.||_2$! For our purposes, a more useful norm is the **operator norm**.

Definition 2.8 (Operator norm)**.** The **operator norm** on the space of linear maps $L(\mathbb{R}^m, \mathbb{R}^n)$ is the value

$$\sup_{x \neq 0} \frac{||L(x)||_2}{||x||_2}. \tag{19}$$

This has the property that it is *sub-multiplicative*, i.e $||AB|| \leq ||A||||B||$ (exercise to reader).

Example 2.9 (2019 P1L)
The matrix function $f(M) = M^{-1}$ is differentiable at I.

Proof. The following proof relies on the fact that the set of $n \times n$ invertible matrices is an open subset in the set of $n \times n$ matrices (see the example sheet).

We first (e.g by checking the one-dimensional case) convince ourselves that the answer is $-I$. This leaves us to verify that

$$\frac{||(I+H)^{-1} - I + H||}{||H||} \to 0. \tag{20}$$

Now we can check that

$$(I+H)^{-1} - I + H = H^2(I+H)^{-1}. \tag{21}$$

So using the sub-multiplicative property, since

$$\frac{||(I+H)^{-1} - I + H||}{||H||} \leq ||H(I+H)^{-1}|| \leq ||H||||(I+H)^{-1}|| \tag{22}$$

we only need show that that $(I+H)^{-1}$ has bounded norm for sufficiently small H[a]. But this is true since suppose $||v||_2 = ||w||_2 = 1$ and

$$(I+H)^{-1}v = \lambda w, \tag{23}$$

Then

$$\frac{1}{\lambda}v = Iw + Hw. \tag{24}$$

Now the magnitude of the RHS is going to arbitrarily close to 1 due to the bounded operator norm of H. So λ can't grow large. □

[a] I can't reason this part as cleanly as I hope the rest of the proof is reasoned!

§3 Notation and Glossary

§3.1 Notation

§3.2 Glossary

- Cts: continuous.

References

[1] Evan Chen (2021), *An Infinitely Large Napkin*, https://venhance.github.io/napkin/Napkin.pdf.