Analysis II ①

1. Uniform Convergence

Suppose we have functions $f_n : E \to \mathbb{R}$ or $\mathbb{C}$, $n \in \mathbb{N}$, ($E \subset \mathbb{R}$ for example). We say that if $\forall x \in E$, $(f_n(x)) \to f(x) \in \mathbb{R}$ then (f_n) converges pointwise or simply to $f : E \to \mathbb{R}$. We would like to infer properties of f from f_n.

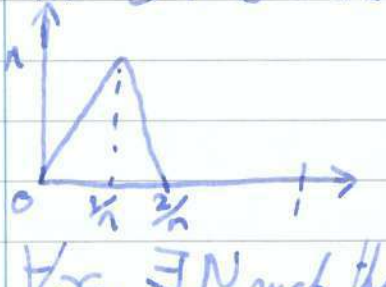
Examples

1. $E = [-1, 1]$, $f_n(x) = x^{\frac{1}{2n+1}}$, $n \in \mathbb{N}_0$

$$f_n(x) \to \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \quad \text{as } n \to \infty$$

So $(f_n) \to f$ pointwise on $[-1, 1]$, and f is not continuous, though each f_n is.

2. $f_n : [0, 1] \to \mathbb{R}$, $f_n(x) = \begin{cases} n^2 x & \\ n^2(\frac{2}{n} - x) & x \in \\ 0 & \end{cases} \begin{matrix} [0, \frac{1}{n}] \\ [\frac{1}{n}, \frac{2}{n}] \\ [\frac{2}{n}, 1] \end{matrix}$



Each f_n is continuous on $[0, 1]$, and ~~$\forall x$~~ $\forall x \in [0, 1]$, $(f_n(x)) \to 0$

$\forall x$, $\exists N$ such that $f_n(x) = 0$ if $n > N$.
So $(f_n) \to 0$ pointwise on $[0, 1]$.

But $\int_0^1 f_n(x)\,dx \not\to 0 = \int_0^1 f(x)\,dx$. To be able to deduce properties of the limiting function, we need a different notion of convergence (there are various).

Definition

Let f_n ($n \in \mathbb{N}$), $f : E \to \mathbb{R}$ (or $\mathbb{C}$) be functions. We say that $(f_n)_{n \geq 0}$ converges uniformly on E to f if:
$\forall \varepsilon > 0$, $\exists N_\varepsilon$ such that $\forall n \geq N_\varepsilon$, $\forall x \in E$, $|f_n(x) - f(x)| < \varepsilon$

Compare this with a definition for pointwise convergence:
$\forall x \in E$, $\forall \varepsilon > 0$, $\exists N_{\varepsilon, x}$ such that $\forall n \geq N$,
$|f_n(x) - f(x)| < \varepsilon$
The difference: "The same n works for every x."

Examples

2. $f_n(x) \geq 1$ if $\frac{1}{n^2} \leq x \leq \left(\frac{2}{n} - \frac{1}{n^2}\right)$. So if $0 < \varepsilon < 1$, then $\forall n$, $\exists x \in [0,1]$ with $|f_n(x)| > \varepsilon$. Therefore, (f_n) does not converge uniformly on $[0,1]$ to 0.

Remark: We say "converges uniformly on E"; the dependence on E is very important.

3. $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = \frac{x}{n} \to 0$ as $n \to \infty$, for every x.
But $\forall \varepsilon > 0$, and $\forall n \in \mathbb{N}$, $\exists x$ such that $|f_n(x)| = |\frac{x}{n}| > \varepsilon$.
So (f_n) does not converge uniformly on $\mathbb{R}$ to 0.

3'. $f_n : [-A, A] \to \mathbb{R}$, $f_n(x) = \frac{x}{n}$
If $\varepsilon > 0$, pick any $N > A/\varepsilon$. Then, $\forall n \geq N$,
$|f_n(x)| = |\frac{x}{n}| < \frac{A}{A/\varepsilon} = \varepsilon$. Therefore (f_n) converges uniformly on $[-A, A]$ to 0.

"Uniform Convergence is not a local property."

Proposition 1

If (f_n) converges uniformly on E to f, then it converges pointwise. So $\forall x \in E$, $f_n(x) \to f(x)$. Then, uniform convergence is stronger than pointwise convergence.

Analysis II ①

The following is another characterisation of uniform convergence :

(Recall that if $g : E \to \mathbb{R}$ is a function, then $\sup_E g$ is the supremum, or least upper bound of $\{g(x) \mid x \in E\}$. Here, E is a non-empty set, and then $\sup_E g$ is a real number, or ∞.

Proposition 2

Let f_n; $f_n : E \to \mathbb{R}$. (f_n) converges uniformly on E to f if and only if $\sup_E |f_n(x) - f(x)| \to 0$ as $n \to \infty$.

f_n "The vertical distance between graphs of f_n and f tends to 0."

Proof

Let $S_n = \sup_E |f_n(x) - f(x)|$. $\forall x \in E$, $|f_n(x) - f(x)| \leq S_n$.
Suppose $S_n \to 0$. Given $\varepsilon > 0$, $\exists N_\varepsilon$ such that $\forall n \geq N_\varepsilon$, then $S_n < \varepsilon$, or equivalently :
$\forall x \in E$, $\forall n > N_\varepsilon$, $|f_n(x) - f(x)| < \varepsilon$. So $(f_n) \to f$ uniformly on E.

Conversely, suppose $(f_n) \to f$ uniformly on E.
Let $\varepsilon > 0$, so that $\exists N_\varepsilon$ such that $\forall n \geq N_\varepsilon$, $\forall x \in E$, $|f_n(x) - f(x)| < \varepsilon$. Then ε is an upper bound for $|f_n(x) - f(x)|$ if $n \geq N_\varepsilon$, so $\varepsilon \geq S_n$.
Therefore, $S_n \to 0$. $\square$

First Application:

Theorem 1

Let $f_n, f : [a,b] \to \mathbb{R}$ be functions such that:
a) Each f_n is continuous on $[a,b]$.
b) $(f_n) \to f$ uniformly on $[a,b]$.
Then, f is continuous on $[a,b]$.

Proof ($\frac{\varepsilon}{3}$ method)

Let $y \in [a,b]$, $\varepsilon > 0$. As $(f_n) \to f$ uniformly on $[a,b]$, $\exists N$ such that $\forall n \geq N$ (although a single n will do), and $\forall x \in [a,b]$, $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$

Also, f_n is continuous, so $\exists \delta > 0$ such that $\forall x \in [a,b]$ with $|x-y| < \delta$, $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$

So $|x-y| < \delta \Rightarrow |f(x) - f(y)| \leq |f_n(x) - f(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$

Then $|f(x) - f(y)| < \varepsilon$, and f is continuous at y. $\square$

Analysis 2 (2)

It is useful to have a criterion for uniform convergence which doesn't involve the limit function.

Also known as "The General Principle of..."

Theorem 2 (Cauchy Criterion for Uniform Convergence)

Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions $f_n : E \to \mathbb{R}$. Then (f_n) is uniformly convergent on E $\iff$ $\forall \varepsilon > 0$, $\exists N = N_\varepsilon$ such that $\forall n, m \geq N$, $\forall x \in E$, $|f_n(x) - f_m(x)| < \varepsilon$

Proof

($\Rightarrow$) Assume $f_n \to f$ uniformly on E. Let $\varepsilon > 0$, then $\exists N$ such that $\forall x \in E$, $\forall n \geq N$, $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$, so $\forall n, m \geq N$, $\forall x \in E$,
$|f_m(x) - f_n(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \varepsilon$

($\Leftarrow$) Assume the criterion holds. By the Cauchy Criterion for sequences in $\mathbb{R}$, the functions f_n converge pointwise to some $f : E \to \mathbb{R}$.
Let $\varepsilon > 0$, and choose N such that $\forall m, n \geq N$, $\forall x \in E$, $|f_m(x) - f_n(x)| < \frac{\varepsilon}{2}$. Fix $m \geq N$.
$|f_m(x) - f(x)| = \lim_{n\to\infty} |f_m(x) - f_n(x)| \leq \frac{\varepsilon}{2} < \varepsilon$
So $(f_n) \to f$ uniformly on E.

Variant

(f_n) is uniformly convergent on E $\iff$ $\sup_{m \geq n,\, x \in E} |f_n(x) - f_m(x)| \to 0$ as $n \to \infty$

Integration

Theorem 3

Suppose f_n $(n \in \mathbb{N})$, f are Riemann integrable on $[a,b]$ with $(f_n) \to f$ uniformly on $[a,b]$. Then $\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx$

Proof

For n sufficiently large, $S_n = \sup_{x\in E} |f(x) - f_n(x)| \in \mathbb{R}$, and $(S_n) \to 0$, so
$\left| \int_a^b f(x) - f_n(x)\,dx \right| \leq \int_a^b |f(x) - f_n(x)|\,dx \leq (b-a) S_n \to 0$

(Soon we will see that f is integrable if $(f_n)_{n\in\mathbb{N}}$ are)

Proposition 3

i) $(f_n) \to f$, $(g_n) \to g$ uniformly on E. Let $a, b \in \mathbb{R}$, then $(af_n + bg_n) \to af + bg$ uniformly on E.

ii) $(f_n) \to f$ uniformly on E, $g : E \to \mathbb{R}$ bounded.
Then $(gf_n) \to gf$ uniformly on E.

N.B. $f_n(x) = \frac{1}{n}$, $g(x) = x$ (not bounded). $f_n \to 0$ uniformly on $\mathbb{R}$, but $gf_n = \frac{x}{n}$ does not converge uniformly on $\mathbb{R}$.

Proof (of (ii))

$\forall x \in E$, $|g(x)| \leq M$. Then $\sup_{x\in E} |f_n g - fg| \leq M \sup_{x\in E} |f_n - f| \to 0$ as $n \to \infty$.

Analysis II ②

Differentiation

This is more subtle as a uniform limit of differentiable functions need not be differentiable.

E.g. $f_n(x) = |x|^{1+\frac{1}{n}}$, $x \in [-1, 1]$, $n \geq 1$. Then:
$\lim_{x \to 0} \frac{f_n(x)}{x} = \text{sgn}(x)\, |x|^{\frac{1}{n}} = 0$, so $f_n'(0)$ exists and $= 0$.
If $f(x) = |x|$, then $(f_n) \to f$ uniformly on $[-1, 1]$, but f is not differentiable at $x = 0$.

In fact, the Weierstrass Approximation Theorem states that any continuous function on $[a, b]$ is a uniform limit of polynomials (so certainly differentiable functions), and there exist continuous functions which are nowhere differentiable.

Theorem 4

Let f_n $(n \in \mathbb{N})$, $g : (a, b) \to \mathbb{R}$; assume that each f_n is continuously differentiable, and $(f_n') \to g$ uniformly on (a, b). Assume, for some $c \in (a, b)$, $(f_n(c))_n$ converges. Then (f_n) converges uniformly on (a, b) and $f(x) = \lim_{n \to \infty} f_n(x)$ is differentiable with derivative g.

(uniform convergence of derivatives implies that limit and derivative can be interchanged, limit of derivative = derivative of limit)

Proof

By proposition 3(i), f_n converges uniformly on (a, b) $\Leftrightarrow$ $(f_n(x) - f_n(c))$ does. As these have the same derivative, we may replace $f_n(x)$ by $f_n(x) - f_n(c)$ i.e. WLOG, we may assume that $f_n(c) = 0 \quad \forall n$.

Then $f_n(x) = \int_c^x f_n'(t)\,dt$ because f_n' is continuous.
Define $F(x) = \int_c^x g(t)\,dt$. g is continuous by Theorem 1, so F is differentiable with $F' = g$.

As $(f_n') \to g$ uniformly on (a,b) (so certainly on any sub-interval):

$$|F(x) - f_n(x)| = \left|\int_c^x g(t) - f_n'(t)\,dt\right| \leq |x-c| \sup_{(a,b)} |g - f_n'|$$
$$\leq (b-a) \sup_{(a,b)} |g - f_n'| \to 0$$
independently of x.

i.e. $(f_n) \to F$ uniformly on (a,b).

Analysis II ③

Series

Given a sequence of functions $g_n : E \to \mathbb{R}$, $n \in \mathbb{N}$, we say the series $\sum g_n$ converges uniformly on E if the sequence of partial sums $f_n = \sum_{k \le n} g_k$ is uniformly convergent on E.
We say $\sum g_n$ converges absolutely uniformly if $\sum |g_n|$ converges uniformly.

Proposition 3

If $\sum g_n$ converges absolutely uniformly on E then it converges uniformly.

Proof

By hypothesis, the series $R_n(x) = \sum_{k=n+1}^{\infty} |g_k(x)|$ converges $\forall x \in E$ and $R_n \to 0$ uniformly on E. So if $r_n(x) = \sum_{k=n+1} g_k(x)$, then $r_n(x)$ is convergent $\forall x \in E$, and since $|r_n(x)| \le R_n(x)$, then $(r_n)_{n \in \mathbb{N}}$ also converges uniformly on E to 0.
i.e. the series $\sum g_n$ is uniformly convergent.

N.B. The series $\sum \frac{(-1)^n x^n}{n}$ converges uniformly on $[0,1)$ and for any $x \in [0,1)$ it converges absolutely, but it does not converge absolutely uniformly on $[0,1)$.

Useful Sufficient Condition "Weierstrass M-Test"

Let $g_n : E \to \mathbb{R}$, $n \in \mathbb{N}$. Suppose $\exists M_n \in \mathbb{R}$, $n \in \mathbb{N}$, and some n_0 such that:

i) $\forall x \in E$, $\forall n \ge n_0$, $|g_n(x)| \le M_n$

ii) $\sum M_n$ is convergent.

Then $\sum g_n$ is absolutely uniformly convergent.

Proof:

Let $\varepsilon > 0$. Then by ii), $\exists N_\varepsilon \geq n_0$ such that $\forall n > m \geq N_\varepsilon$, $M_m + \dots + M_n < \varepsilon$. Let $f_n = \sum_{k=1}^{n} |g_k|$ be the n^{th} partial sum of $\sum |g_k|$. Then, $\forall x \in E$, and $n > m \geq N_\varepsilon$, $|f_m(x) - f_n(x)| = |g_{m+1}(x)| + \dots + |g_n(x)| < \varepsilon$, by i).

By the Cauchy Criterion, f_n is uniformly convergent on E, so $\sum g_n$ is absolutely uniformly convergent on E.

Power Series: Theorem 5

Let $\sum_{n=0}^{\infty} c_n (x-a)^n$ be a (real or complex) power series. Then, $\exists! R \in [0, \infty]$ called the radius of convergence of the series such that:

a) $|x-a| < R \Rightarrow \sum c_n (x-a)^n$ converges absolutely

b) $|x-a| > R \Rightarrow$ Divergence.

c) Moreover, if $0 \leq r < R$ then $\sum c_n (x-a)^n$ converges absolutely uniformly on $\{x \mid |x-a| \leq r\}$. (If $R = \infty$ then this holds for any $r \in \mathbb{R}_{\geq 0}$)

Proof:

The existence and uniqueness of R from Analysis I covers a) and b). $R = \sup\{r \in [0, \infty) \mid \forall x \text{ with } |x-a| \leq r, \sum c_n (x-a)^n \text{ converges}\}$

c). Let $0 \leq r < R$, and choose r_1 with $r < r_1 < R$. Then, by a), $\sum |c_n| r_1^n$ is convergent, so in particular, $\exists B$ such that $\forall n$,

$|c_n| r^n \leq B$.

Then $|x-a| \leq r \Rightarrow |c_n(x-a)^n| \leq |c_n| r^n \leq B\left(\frac{r}{r_1}\right)^n \leq M_n$

The geometric series $\sum M_n$ is convergent since $r < r_1$.

Therefore, $\sum c_n (x-a)^n$ is absolutely uniformly convergent for $|x-a| \leq r$. □

Remarks

1. Part c) can be phrased as "a power series is locally uniformly convergent" within the area $|x-a| \leq R$ of convergence".

2. The series $\sum_{n=0}^{\infty} x^n$ is not uniformly convergent on $(-1, 1)$, but it is uniformly convergent on $[-r, r]$ for $r < 1$.

Theorem 6

Let $\sum_{n=0}^{\infty} c_n (x-a)^n$ be a power series with radius of convergence $R > 0$ representing the function $f(x)$, $x \in \{|x-a| < R\}$. Then:

a) The "derived series" $g(x) = \sum_{n=1}^{\infty} n c_n (x-a)^{n-1}$ also has radius of convergence R.

b) f is differentiable on $\{x \mid |x-a| < R\}$ with $f' = g$.

Proof (WLOG assume $a = 0$)

a) Let R_1 be the radius of convergence of g. Now, $\forall n \geq 1$, $|c_n x^n| \leq |x| |n c_n x^{n-1}|$. So if the series for $g(x)$ converges absolutely, so does the series for $f(x)$, hence $R_1 \leq R$. If $R_1 < R$, we could choose r_1, r with $R_1 < r_1 < r < R$, then $\sum n |c_n| r_1^{n-1}$ diverges, and $\sum |c_n| r^n$ converges. This is impossible since $\frac{n r_1^{n-1}}{r^n} \to 0$ as $n \to \infty$.

b) Let $f_n(x) = \sum_{k=0}^{n} c_k x^k$. Then, for any $r < R$, the sequence $(f_n'(x))_{n\in\mathbb{N}}$ converges uniformly for $|x| \leq r$ to $g(x)$, by a) and Theorem 5 c).

So we can apply Theorem 4 to see that f is differentiable for $|x| < R$ with derivative g. So this holds for any x with $|x| < R$ (simply choosing $|x| < r < R$).

Analysis II ④

Uniform Continuity and Integration

Definition

Let $I \subset \mathbb{R}$, an interval, and $f : I \to \mathbb{R}$ is *uniformly continuous* on I if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\forall x, y \in I$ with $|x - y| < \delta$, we have $|f(x) - f(y)| < \varepsilon$

The difference between simple and uniform continuity is that δ is independent of x.

a) $f : (0, 1] \to \mathbb{R}$, $f(x) = \frac{1}{x}$
If $\delta > 0$, $0 < x < \delta$, $y = \frac{x}{2}$ then $|x - y| < \delta$, but
$|f(x) - f(y)| = |\frac{1}{x} - \frac{2}{x}| = \frac{1}{x} \geq 1$
So f is not uniformly continuous on $(0, 1]$.

b) $f : (0, 1) \to \mathbb{R}$. $f(x) = \sin(1/x)$, bounded.
$\forall n \geq 1$, $f(\frac{1}{n\pi}) = 0$, $f(\frac{1}{(n+\frac{1}{2})\pi}) = (-1)^n$. For any $\delta > 0$, choose n such that $n(2n+1) > \frac{1}{\pi\delta}$
If $x = \frac{1}{n\pi}$, $y = \frac{1}{(n+\frac{1}{2})\pi}$, then $|x - y| < \delta$, but $|f(x) - f(y)| = 1$.
Therefore, f is not uniformly convergent on $(0, 1)$.

Theorem 7

(c o f u c d o b i i u c!

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous on $[a, b]$. (We can replace $[a, b]$ with any compact metric space and $\mathbb{R}$ by any metric space.)

Proof

Suppose not. Then, $\exists \varepsilon > 0$ such that $\forall n > 0$, $\exists x_n, y_n \in [a, b]$ with $|x_n - y_n| < \frac{1}{n}$, $|f(x_n) - f(y_n)| \geq \varepsilon$.

$(x_n)_{n\in\mathbb{N}}$ is bounded, so by Bolzano-Weierstrass, contains a convergent subsequence $x_{\sigma(n)}$ (where $\sigma : \mathbb{N} \to \mathbb{N}$ is an increasing map), and $|x_{\sigma(n)} - y_{\sigma(n)}| < \frac{1}{\sigma(n)} \leq \frac{1}{n}$

So by replacing (x_n, y_n) with $(x_{\sigma(n)}, y_{\sigma(n)})$, we may assume WLOG that $(x_n) \to x \in [a,b]$, $y_n = x_n - (x_n - y_n)$ $\Rightarrow y_n \to x$, as $n \to \infty$, since $x_n - y_n \to 0$. So, as f is continuous, $f(x_n) \to f(x)$, $f(y_n) \to f(x)$. But we assumed that $\forall n$, $|f(x_n) - f(y_n)| \geq \varepsilon$ $\#$

1st Application: Riemann Integration

Recall $f : [a,b] \to \mathbb{R}$, bounded. Consider "dissections":
$\mathcal{D} = (a = a_0 < a_1 < \dots < a_N = b)$ and upper and lower sums:
$S(f, \mathcal{D}) = \sum_{j=0}^{N-1} (a_{j+1} - a_j) \sup_{[a_j, a_{j+1}]} f$

$s(f, \mathcal{D}) = \sum_{j=0}^{N-1} (a_{j+1} - a_j) \inf_{[a_j, a_{j+1}]} f$. f bounded $\Rightarrow$ sup, inf exist

$I^*(f) = \inf_{\mathcal{D}} S(f, \mathcal{D}) \geq \sup_{\mathcal{D}} s(f, \mathcal{D}) = I_*(f)$

We say that f is Riemann Integrable if $I^*(f) = I_*(f) = \int_a^b f(x)\,dx$

Theorem: Riemann's Criterion

f is integrable $\iff \forall \varepsilon > 0, \exists \mathcal{D}$ such that $S(f, \mathcal{D}) - s(f, \mathcal{D}) < \varepsilon$.

Theorem 8

Let $f : [a,b] \to [A,B] \subset \mathbb{R}$ be integrable, and $g : [a,b] \to \mathbb{R}$ continuous. Then $g \circ f : [a,b] \to \mathbb{R}$ is integrable.

Analysis II ④

Corollary

$g : [a, b] \to \mathbb{R}$ continuous $\Rightarrow g$ integrable (take $f(x) = x$)

Proof:

Let $\varepsilon > 0$. By Theorem 7, g is uniformly continuous on $[a, b]$. So $\exists \delta > 0$ such that $\forall x, y \in [a, b]$, with $|x - y| < \delta$, we have $|g(x) - g(y)| < \varepsilon$. WLOG, we may assume that $\delta < \varepsilon$, (since any smaller δ value will work).

Choose $D = (a = a_0 < a_1 < \ldots < a_n = b)$ such that $S(f, D) - s(f, D) < \delta^2$

Write U_j, u_j = sup, inf of f on $[a_j, a_{j+1}]$

W_j, w_j = sup, inf of g on $[u_j, U_{j+1}]$

Let $J = \{ j \mid U_j - u_j < \delta \}$ "good guys"

Then by choice of δ, if $j \in J$, $W_j - w_j \leq \varepsilon$.

Also, $\sum_{j \notin J} \delta (a_{j+1} - a_j) \leq \sum_{j \notin J} (U_j - u_j)(a_{j+1} - a_j) \leq \sum (U_j - u_j)(a_{j+1} - a_j)$ "bad guys"

$= S(f, D) - s(f, D) < \delta^2$

i.e. $\sum_{j \notin J} (a_{j+1} - a_j) < \delta$

$\therefore S(g \circ f, D) - s(g \circ f, D) = \sum_j (W_j - w_j)(a_{j+1} - a_j)$

$= \sum_{j \in J} \square + \sum_{j \notin J} \square \leq \underset{(j \in J)}{\varepsilon (b - a)} + \underset{(j \notin J)}{2 \sup_{[a,b]} |g| \, (a_{j+1} - a_j)}$

$\leq \varepsilon (b - a + 2 \sup_{[a,b]} |g|)$

Then by Riemann's Criterion, $g \circ f$ is integrable. $\square$

Analysis II (5)

Proposition 4

If $f_n : [a,b] \to \mathbb{R}$, $n \in \mathbb{N}$, are ~~d~~ integrable, and $(f_n) \to f$ uniformly, then f is also integrable.

Proof:

Let $c_n = \sup_{[a,b]} |f - f_n|$, so that $c_n \to 0$.
So $\forall x \in [a,b]$, $f_n(x) - c_n \leq f(x) \leq f_n(x) + c_n$
So $\forall n, \mathcal{D}$, $s(f_n, \mathcal{D}) - (b-a)c_n = s(f_n - c_n, \mathcal{D})$,
$S(f_n, \mathcal{D}) + (b-a)c_n = S(f_n + c_n, \mathcal{D})$

$s(f_n - c_n, \mathcal{D}) \leq s(f, \mathcal{D}) \leq S(f, \mathcal{D}) \leq S(f_n + c_n, \mathcal{D})$
As f_n is integrable, and $c_n \to 0$, then $\forall \varepsilon > 0$ $\exists n$ and $\mathcal{D}$ such that $(b-a)c_n < \frac{\varepsilon}{3}$, and $S(f_n, \mathcal{D}) - s(f_n, \mathcal{D}) < \frac{\varepsilon}{3}$
$\Rightarrow S(f, \mathcal{D}) - s(f, \mathcal{D}) < \varepsilon$, so by Riemann's Criterion, this is integrable.

Now we extend integration to complex and vector valued functions.

Definition

$f = (f_1, \dots, f_n) : [a,b] \to \mathbb{R}^n$. If the components, f_i, are all (Riemann) integrable on $[a,b]$, we say that f is Riemann Integrable and define:
$$\int_a^b f(x)\,dx = \left(\int_a^b f_1\,dx, \dots, \int_a^b f_n\,dx\right) \in \mathbb{R}^n$$

Suppose $f : [a,b] \to \mathbb{C}$ with $Re(f), Im(f) : [a,b] \to \mathbb{R}$ both integrable. Then we say that f is integrable and:
$$\int_a^b f(x)\,dx = \int_a^b Re[f(x)]\,dx + i\int_a^b Im[f(x)]\,dx$$

Remarks

1. Integration of functions $\mathbb{C} \to \mathbb{C}$ or $\mathbb{R}^n \to \mathbb{R}^n$ is a different matter.

2. If $A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, and $f : [a,b] \to \mathbb{R}^n$ is integrable, then so is Af (by linearity of integration). So integration is basis independent (i.e. it makes sense to talk of $\int f$, where $f : [a,b] \to V$, V a finite dimensional vector space).

If $f : [a,b] \to \mathbb{R}$ is integrable, then we know $\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$. This is similarly true for vector valued integrals. Write $\|(x_1, \dots, x_n)\| = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$, Euclidean norm or distance

Proposition 5

Let $F : [a,b] \to \mathbb{R}^n$ be integrable. Then $\|F\| : [a,b] \to \mathbb{R}$ is also integrable, and:

$$\left\|\int_a^b F(x)\,dx\right\| \leq \int_a^b \|F(x)\|\,dx$$

Proof

$\|F(x)\| = \left(\sum_{i=1}^n f_i(x)^2\right)^{\frac{1}{2}}$. As $x \mapsto x^2$, $x \mapsto \sqrt{x}$ $(x \geq 0)$ are continuous, and sums of integrable functions are integrable, Theorem 8 says that $\|F\|$ is integrable.

Let $v = \int_a^b F(x)\,dx \in \mathbb{R}^n$. If $v = 0$, we have nothing to prove. Otherwise:

$$\|v\|^2 = \sum_{i=1}^n v_i^2 = \sum_{i=1}^n \int_a^b v_i f_i(x)\,dx = \int_a^b \left(\sum_{i=1}^n v_i f_i(x)\right) dx \leq \int_a^b \|v\|\,\|F(x)\|\,dx$$

by Cauchy-Schwartz. Now simply divide by $\|v\|$, and remember $\|v\| = \left\|\int_a^b F\,dx\right\|$

Analysis II ⑤

Particular Case

$f : [a, b] \to \mathbb{C} \cong \mathbb{R}^2$, $x + iy \mapsto (x, y)$
$\|(x, y)\| = |x + iy|$
So f integrable $\Rightarrow |f|$ integrable, and $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$

To illustrate the theory:

Theorem 9 (Weierstrass Approximation Theorem)

Let $f : [0, 1] \to \mathbb{R}$ (or $\mathbb{C}$) be continuous. Then, f is the uniform limit of a sequence of polynomials on $[0,1]$. In fact, the sequence
$f_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$, $n \geq 1$
(Bernstein Polynomials) converge uniformly to f on $[0, 1]$.

Of course, there are many sequences of polynomials converging uniformly to f (namely $f_n + g_n$, where $g_n \to 0$ uniformly).

Proof

Let $p_{k,n}(x) = \binom{n}{k} x^k (1-x)^{n-k}$
Then, $\sum_{k=0}^{n} p_{k,n}(x) = 1$ by the Binomial Theorem ①

Therefore, $\forall x \in [0, 1]$, $0 \leq p_{k,n}(x) \leq 1$.
$k\, p_{k,n}(x) = nx\, p_{k-1,n-1}(x)$
$k(k-1)\, p_{k,n}(x) = n(n-1)x^2\, p_{k-2,n-2}(x)$

$\sum_{k=0}^{n} k\, p_{k,n}(x) = nx$ by ①

$\sum_{k=0}^{n} k^2 p_{k,n}(x) = \sum_{k=0}^{n} (k + k(k-1))\, p_{k,n}(x) = nx + n(n-1)x^2$
$= n^2x^2 + nx(1-x)$

$\Rightarrow \sum_{n=0}^{k} (nx-k)^2 p_{k,n}(x) = nx(1-x)$ ②

3

if we replace x by $f(x)$

(i.e. $\sum_{k=0}^{n} \frac{k}{n} p_{k,n}(x) = x$, $= f_n$ if $f(x) = x$)

Let $\varepsilon > 0$. As f is uniformly continuous on $[0,1]$, $\exists \delta > 0$ such that $\forall x, y \in [0,1]$ with $|x-y| < \delta$, we have $|f(x)-f(y)| < \varepsilon/2$

Now $|f(x) - f_n(x)| = \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) p_{k,n}(x) \right|$ by ①

$$\leq \left| \sum_{\substack{k=0 \\ |x-k/n| < \delta}}^{n} \square \right| + \left| \sum_{\substack{k=0 \\ |x-k/n| \geq \delta}}^{n} \square \right| = A + B \text{ say.}$$

$A < \sum_{k, |x-k/n| < \delta} \frac{\varepsilon}{2} p_{k,n}(x) \leq \frac{\varepsilon}{2}$ by ①

$B \leq \sum_{k, |x-k/n| \geq \delta} 2M p_{k,n}(x)$ where $M = \sup_{[a,b]} |f(x)|$

$\leq \frac{2M}{\delta^2} \sum_{k, |x-k/n| \geq \delta} \left(x - \frac{k}{n}\right)^2 p_{n,k}(x) \leq \frac{2M}{\delta^2 n^2} (nx(1-x))$ by ②

$\leq \frac{2M}{n\delta^2} < \frac{\varepsilon}{2}$ provided $n > \frac{4M}{\delta^2 \varepsilon}$.

So $|f(x) - f_n(x)| < \varepsilon$ for such n.

$\Rightarrow (f_n) \to f$ uniformly on $[0,1]$ $\square$

Analysis II ⑥

$\mathbb{R}^n$ as a Normed Space

We study functions $f: \mathbb{R}^n \to \mathbb{R}^m$. The role of $|x-y|$ in $\mathbb{R}$ is played by the Euclidean distance in $\mathbb{R}^n$. It is both useful and convenient to generalise to the idea of a norm

Definition

Let V be a vector space (over $\mathbb{R}$ or $\mathbb{C}$). A norm on V is a function $\|\cdot\|: V \to \mathbb{R}$ satisfying:

1. $\forall v \in V$, $\|v\| \geq 0$, and $\|v\| = 0 \Leftrightarrow v = 0$
2. $\forall v \in V$, $\forall a \in \mathbb{R}$ (or $\mathbb{C}$), $\|av\| = |a| \, \|v\|$
3. $\forall v, w \in V$, $\|v+w\| \leq \|v\| + \|w\|$ ("triangle inequality")

A normed space is a pair $(V, \|\cdot\|)$, where $\|\cdot\|$ is a norm on V

Remark

If it is clear what the norm is, we may just say that "V is a normed space", but in general, a space can have many different norms.

Example

$V = \mathbb{R}^n$

- Euclidean norm $\|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$, $x \in \mathbb{R}^n$
- $\mathbb{C}$, $\|z\| = |z|$, or $\mathbb{C}^n$, $\|z\| = \left(\sum_{i=1}^{n} |z_i|^2\right)^{\frac{1}{2}}$, $z \in \mathbb{C}^n$

- $\mathbb{R}^n$, $\|v\|_\infty = \max\{|v_1|, \dots, |v_n|\}$

- $\mathbb{R}^n$, $\|v\|_1 = |v_1| + \dots + |v_n|$, the "taxicab norm"

The only non-trivial axiom is 3. For example, for $(\mathbb{R}^n, \|\cdot\|_2)$,

$$\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2|v\cdot w| \leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\|$$
$$= (\|v\| + \|w\|)^2 \qquad \text{(by Cauchy-Schwarz)}$$

Other norms on $\mathbb{R}^n$: "L^p norm", $p \in \mathbb{R}$, $p \geq 1$.
$\|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}$ (proof that it satisfies 3 uses Hölder's inequality).

Sequence Spaces

$\mathbb{R}^{\mathbb{N}}$ is the space of all sequences $(x_k)_{k\in\mathbb{N}}$ of reals under termwise addition and termwise scalar multiplication $\alpha(x_k) = (\alpha x_k)$.

$L_\infty = \{\text{bounded sequences}\} = \{(x_k) \in \mathbb{R}^{\mathbb{N}} \mid \exists R \text{ such that } \forall k, |x_k| < R\}$
with norm $\|(x_k)\|_\infty = \sup_k \{|x_k|\}$

$L_2 = \{(x_k) \in \mathbb{R}^{\mathbb{N}} \mid \sum_k x_k^2 \text{ is convergent}\}$ with norm
$\|x\|_2 = \left(\sum x_k^2\right)^{\frac{1}{2}}$.
Notice that $\|(x_k)\| = \lim_{n\to\infty} \|(x_0, x_1, \dots, x_{n-1})\|$, so the triangle inequality follows.

Analysis II (6)

Function Spaces

$C[a,b] = \{\text{continuous functions } f : [a,b] \to \mathbb{R}\}$
This is a vector space under pointwise addition and multiplication.

- "sup-norm", "uniform norm", $\|f\|_\infty = \sup_{[a,b]} \{|f(x)|\}$
- "L_2-norm", $\|f\|_2 = \left(\int_a^b |f(x)|^2 \, dx \right)^{\frac{1}{2}}$

If $\|\cdot\|$ is a norm on V, then so is $c\|\cdot\|$ for any $c > 0$.
If, say, $c > 1$, and we have another norm $\|\cdot\|'$ with $\|v\| < \|v\|' < c\|v\|$, then $\|\cdot\|'$ will behave much like $\|\cdot\|$

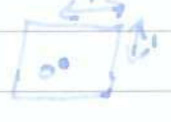
Definition

Let V be a vector space and $\|\cdot\|$, $\|\cdot\|'$ norms on V. They are Lipschitz Equivalent if $\exists a, b \in \mathbb{R}$ with $0 < a < b$, such that $\forall v \in V$, $a\|v\| < \|v\|' \leq b\|v\|$. This is an equivalence relation on norms.

Reformulation

Define the "open ball of radius r" $B_r(a) = \{v \in V \mid \|v - a\| < r\}$.

For $\mathbb{R}^2$, $\| \ \|_2$, $B_0(r)$ is 

For $\| \ \|_\infty$, we have 

Proposition 6

$\|\cdot\|$, $\|\cdot\|'$ are equivalent $\iff \exists r, R \in \mathbb{R}$ with $0 < r < R$ such that $B_0(r) \subset B'(0,1) \subset B(0,R)$

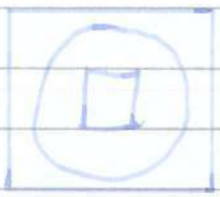
Proof

Suppose that $\|\cdot\|$, $\|\cdot\|'$ are equivalent, so $a\|v\| \leq \|v\|' \leq b\|v\|$.
Then $v \in B(\frac{1}{b}) \Rightarrow \|v\|' \leq b\|v\| < b\frac{1}{b} \Rightarrow B(\frac{1}{b}) \subset B'(1)$
Similarly, $B'(1) \subset B(\frac{1}{a})$

Conversely, suppose $B'(1) \subset B(R)$. Let $0 \neq v \in V$, and $c = \frac{1}{2\|v\|'}$
Then $\|cv\|' = \frac{1}{2\|v\|'}\|v\|' = \frac{1}{2} < 1$, $\Rightarrow cv \in B'(1)$, $cv \in B(R)$
The other inequality is gained similarly.

Example

On $\mathbb{R}^n$, $\|\cdot\|_{Eucl}$ and $\|\cdot\|_\infty$ are equivalent. Remark: later we will show that any two norms on $\mathbb{R}^n$ are equivalent.



Analysis II ⑦

Example of Inequivalent Norms

$l_2 = \{ \text{sequences } (x_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum x_n^2 < \infty \}$
$l_\infty = \{ \text{bounded sequences in } \mathbb{R} \}$

If $(x_n) \in l_2$, then $(x_n) \to 0$, so in particular $(x_n) \in l_\infty$
$\therefore$ On l_2 we have two norms.

$\|x\| = \|x\|_2$, $\|x\|' = \|x\|_\infty$
Clearly $\|\cdot\| \geq \|\cdot\|'$. These norms are not equivalent.

Let $v^{(n)} = (\underbrace{\tfrac{1}{2}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}}_{n \text{ times}}, 0, 0, \ldots) \in l_2$, $\|v^{(n)}\| = \frac{\sqrt{n}}{2}$
$\|v^{(n)}\|' = \frac{1}{2}$. So $\{v^{(n)} \mid n \in \mathbb{N}\} \subset B'(0,1)$ (l_∞ norm)
but $\nexists R$ such that $\{v^{(n)} \mid n \in \mathbb{N}\} \subset B(0,R)$ (l_2 norm).
i.e. $B'(0,1) \not\subset B(0,R)$ for any $R > 0$, so the norms are not equivalent, by proposition 6.

Sequences and Convergence

Definitions

Let $(V, \|\cdot\|)$ be a normed space.

i) A subset $E \subset V$ is bounded if $\exists R$ such that $\forall x \in E$, $\|x\| \leq R$

ii) If x_k $(k \in \mathbb{N})$, and $x \in V$, we say $(x_k) \to x$ if $\|x_k - x\| \to$
Equivalently, $\forall \varepsilon > 0$, $\exists N$ such that $k \geq N \Rightarrow \|x_k - x\| < \varepsilon$.

$\mathbb{R}^n$ with Euclidean Norm

For a sequence $x^{(k)} \in \mathbb{R}^n$, $x^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$.

Proposition 7

Let $x^{(k)}, x \in \mathbb{R}^n$. Then $(x^{(k)}) \to x$ (for $\|\cdot\|_2$)
$\iff \forall i \in \{1, \dots, n\}, \; x_i^{(k)} \to x_i$.

Proof:

$(x^{(k)}) \to x \iff \sum_{i=1}^{n} (x_i - x_i^{(k)})^2 \iff \forall i, \; x_i^{(k)} \to x_i$

Important Example

Consider $C[a,b]$ with norm $\|f\| = \sup_{[a,b]} |f|$.
f_n $(n \in \mathbb{N})$, $f \in C[a,b]$. By proposition 7,
$\sup_{[a,b]} |f_n - f| \to 0 \iff (f_n) \to f$ uniformly on $[a,b]$.
So convergence in $(C[a,b], \|\;\|_{sup})$ is just uniform convergence.

Facts about convergence in a normed space

- $(x_k) \to x$, $(x_k) \to y \Rightarrow x = y$
(since $\|x - y\| \leq \|x - x_k\| + \|y - x_k\| \to 0$ as $k \to \infty \Rightarrow x = y$)

- $(x_k) \to x$, $(y_k) \to y \Rightarrow (ax_k) \to ax$, $a \in \mathbb{R}$
$(x_k + y_k) \to x + y$

N.B. the definition of convergence depends on $\|\cdot\|$, not just V.

Proposition 8

Suppose $\|\cdot\|$, $\|\cdot\|'$ are equivalent norms on V. Then:

i) $E \subset V$ is bounded in $(V, \|\cdot\|) \iff E \subset V$ is bounded in $(V, \|\cdot\|')$

ii) (x_k) in V is convergent in $(V, \|\cdot\|) \iff$ it is convergent in $(V, \|\cdot\|')$

Analysis II ⑦

Proof:

i) This follows from the definition of equivalence.

ii) Suppose that $\|v\|' \leq b\|v\|$, $\forall v \in V$ ①
Then (x_k) is convergent in $(V, \|\cdot\|) \Leftrightarrow \exists x \in V$ such that $\|x_k - x\| \to 0$. But ① then implies (for the same x) that $\|x_k - x\|' \to 0$ i.e. $(x_k) \to x$ in $(V, \|\cdot\|')$. The other direction follows by symmetry.

Later we will see why this is true : $id : (V, \|\cdot\|) \to (V, \|\cdot\|')$ and its inverse are continuous).

Theorem 10 (Bolzano-Weierstrass Theorem for $\mathbb{R}^n$)

Any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence (N.B $(\mathbb{R}^n, \|\cdot\|_{Eucl})$.

Proof:

a) Lion-hunting (n dimensional boxes instead of intervals)

b) Induction on n. For $n=1$, this is the usual Bolzano-Weierstrass Theorem in $\mathbb{R}$.

Assume this holds in $\mathbb{R}^{n-1}$. Let $(x^{(k)})$ be a bounded sequence in $\mathbb{R}^n$, say $\|x^{(k)}\| \leq R$.

Let $y^{(k)} = (x_1^{(k)}, \dots, x_{n-1}^{(k)}) \in \mathbb{R}^{n-1}$
By our induction hypothesis, $(y^{(k)})$ contains a convergent subsequence $(y^{\sigma(k)}) \to y \in \mathbb{R}^{n-1}$, where $\sigma : \mathbb{N} \to \mathbb{N}$ is a strictly increasing map.

Now the sequence $(x_n^{\sigma(k)})$ in $\mathbb{R}$ is bounded and so contains a convergent subsequence, $(x_n^{(\sigma'(k))}) \to y_n \in \mathbb{R}$.
Then $\forall i$, $x_i^{\sigma'(k)} \to y_i \Rightarrow x \to y$, by Proposition 7.

Important Remark

This doesn't immediately generalise to normed spaces.

Example

$l_\infty = \{$bounded sequences in $\mathbb{R}^n\}$. $\|x\|_\infty = \sup_n |x_n|$.
Let $e_k = (0, 0, \ldots, 0, 1, 0, \ldots) \in l_\infty$, with a 1 in place k.
$\|e_k\| = 1$ so (e_k) is bounded.

We claim that this has no convergent subsequence. If it did, say $(e_{\sigma(k)}) \to x$, then let $\varepsilon > 0$, and $\exists N$ such that $\forall k \geq N$,
$\|e_{\sigma(k)} - x\| < \varepsilon$
$\Rightarrow \|e_{\sigma(k)} - e_{\sigma(k+1)}\| < \|e_{\sigma(k)} - x\| + \|e_{\sigma(k+1)} - x\| < 2\varepsilon$

But $\|e_{\sigma(k)} - e_{\sigma(k+1)}\| = 1$. So we have a contradiction when $\varepsilon < \frac{1}{2}$.

Analysis II (8)

Cauchy Sequences and Completeness

Definition

Let $(V, \|\cdot\|)$ be a normed space. We say a sequence (x_n) in V is a Cauchy Sequence if $\forall \varepsilon > 0$, $\exists N = N_\varepsilon$ such that $\forall m, n \geq N$, $\|x_m - x_n\| < \varepsilon$. For $V = \mathbb{R}$, this is just the usual.

Proposition 9

i) If $(x_n) \to x$ converges, then x_n is a Cauchy sequence.

ii) A Cauchy sequence is bounded.

Proof

i) Given $\varepsilon > 0$, $\exists N$ such that $\forall m \geq N$, $\|x_m - x\| < \frac{\varepsilon}{2}$, since $(x_n) \to x$. $\forall m, n \geq N$, $\|x_m - x_n\| \leq \|x_m - x\| + \|x_n - x\| < \varepsilon$.

ii) $n \geq N_\varepsilon$: $\|x_n\| \leq \|x_n - x_N\| + \|x_N\| < \varepsilon + \|x_N\|$
so $\forall n \in \mathbb{N}$, $\|x_n\| < \varepsilon + \max\{\|x_i\| \mid i \leq N\}$

In general, Cauchy sequences needn't converge.

Definition

$(V, \|\cdot\|)$ is complete if every Cauchy sequence is convergent in V.
(N.B. these definitions all depend on V and $\|\cdot\|$, not just V)

Theorem 11

$\mathbb{R}^n$ (with Euclidean norm) is complete.

Proof:

Let $(x^{(k)})$ be a Cauchy Sequence in $\mathbb{R}^n$. Then, $\forall \varepsilon$, $\exists N_\varepsilon$ such that $\forall j, k \geq N_\varepsilon$, $\|x^{(j)} - x^{(k)}\| < \varepsilon$, so for $1 \leq i \leq n$, $|x_i^{(j)} - x_i^{(k)}| < \varepsilon$.

So each sequence $(x_i^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, so converges to some $x_i \in \mathbb{R}$. Then by Proposition 7, $(x^{(k)}) \to x$.

Theorem 12

$C[a,b]$ with uniform-norm is complete.

Proof:

Combine Theorem 2 (Cauchy criterion for uniform convergence) and Theorem 1 (Uniform limits of continuous functions are continuous)

Examples of Normed Spaces which are not complete

1. $L_0 = \{(x_n) \in \mathbb{R}^{\mathbb{N}} \mid x_n = 0$ for all but finitely many $n\}$ with norm $\|(x_n)\|_\infty = \sup\{|x_n|\}$. L_0 is a subspace of L_∞. (Exercise: L_∞ is complete with respect to $\|\cdot\|_\infty$).

Let $x^{(k)} = (1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{k}, 0, 0, \dots) \in L_0$, $\forall k \geq 1$

$\|x^{(k)} - x\|_\infty = \sup\{\frac{1}{k+1}, \frac{1}{k+2}, \dots\} = \frac{1}{k+1} \to 0$. So $(x^{(k)})$ is a Cauchy sequence in L_∞ (and also in $(L_0, \|\cdot\|_\infty)$). In $(L_\infty, \|\cdot\|_\infty)$, $x^{(k)} \to (1, \frac{1}{2}, \frac{1}{3}, \dots)$

2.

But $(x^{(k)})$ does not converge in ℓ_0, for if $(x^{(k)}) \to y \in \ell_0$, then since $\ell_0 \subset \ell_\infty$, $(x^{(k)}) \to y$ in ℓ_∞. But then $y = x$ by uniqueness of limits, and $x \notin \ell_0$.

2. $V = \mathbb{R}[x]$, all real polynomials, with norm $\|p(x)\| = \sup_{[0,1]} |p(x)|$
N.B. This is a norm because if $p(x) = 0$, $\forall x \in [0,1]$, then $p(x) \equiv 0$.

Let $p_n = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}$. Then $(p_n(x)) \to e^x$ uniformly on $[0,1]$ (by the M-test) $\Rightarrow (p_n)$ is a Cauchy sequence for the uniform norm on $[0,1]$, but doesn't converge in $\mathbb{R}[x]$ (It does in $C[0,1]$).

Topology of Normed Spaces

Basic Definitions.

Let $(V, \|\cdot\|)$ be a normed space, and $E \subseteq V$.

- We say that $x \in V$ is a limit point of E if $\exists (x_k) \in E$ with all $x_k \neq x$, and $(x_k) \to x$.
$V = \mathbb{R}$, $E = (0,1) \Rightarrow$ Both $0, 1$ are limit points of E, because $(\frac{1}{n})_{n \geq 2} \to 0$, $(1 - \frac{1}{n})_{n \geq 2} \to 1$.
(Also, every $x \in (0,1)$ is a limit point of $(0,1)$: $(x + \frac{1}{n})_{n \geq n_0} \to x$, for $n_0 > \frac{1}{1-x}$).
But if $E = \{0, 1, 2\} \subset V = \mathbb{R}$, then $0, 1$ are not limit points of E, because if $(x_k) \to 0$, then $x_k = 0$ for all sufficiently large k.

- A subset $E \subset V$ is closed if every limit point of E belongs to E.
$(0,1)$ is not closed in this sense. $E = \{0, 1, 2\}$ is closed.

- $E \subset V$ is open if $\forall x \in E$, $\exists r > 0$ such that $B(x, r) \subset E$.

Theorem 13

i) $E \subset V$ is open $\Leftrightarrow$ Its complement $V \setminus E$ is closed.

ii) Any union of open sets is open. Any intersection of a finite collection of open sets is open.

iii) V, $\emptyset$ are open and closed.

iv) Every ball $B(x, r) \subset V$ is open.

N.B. The notions of a subset being open (or closed) are relative to the space V.

Proposition 14

x is a limit point of $E \subset V \Leftrightarrow \forall \varepsilon > 0$, $\exists y \in E$ with $y \neq x$, $\|x - y\| < \varepsilon$.

Proof

($\Rightarrow$) If $(x_k) \to x$, $x_k \in E \setminus \{x\}$, then $\forall \varepsilon > 0$, $\exists k$ with $\|x_k - x\| < \varepsilon$, and take $y = x_k$.

($\Leftarrow$) Take $\varepsilon = 1/n$. Choose $x_n \neq x$ in E with $\|x_n - x\| < \varepsilon$. Then $(x_n) \to x$.

Analysis II (9)

i) E open $\Leftrightarrow V \setminus E$ closed

Suppose E is open and let $F = V \setminus E$, and $x \in V$ be a limit point of F. We must show that $x \in F$. If not, then $x \in E$ so $\exists r > 0$ such that $B(x, r) \subset E$. Then $\forall y \in F$, $\|y - x\| \geq r$ contradicting Proposition 14 (since x is a limit point).

Conversely, let F be closed, and $x \in E$, not a limit point of F. So by proposition 14, $\exists r > 0$ such that $\forall y \in F$, $\|x - y\| \geq r \Rightarrow B(x, r) \subset E$, E is open.

ii) $\bigcup_{i \in I} U_i$ is open, and $\bigcap_{i=1}^{n} U_i$ is open, for U_i open

This follows for unions from the definition of open. For intersections, let $U_1, \ldots, U_n \subset V$ be open and $x \in \bigcap U_i$. Then for each i, $\exists r_i > 0$ such that $B(x, r_i) \subset U_i$. Let $r = \min\{r_i\} > 0$. Then $B(x, r) \subset \bigcap_{i=1}^{n} U_i$, so $\bigcap_{i=1}^{n} U_i$ is open.

iii) V and $\emptyset$ are both open by definition, so they are also both closed.

iv) To show that $B(a, R) \subset V$ is open, let $x \in B(a, R)$ and let $r = \|x - a\|$, $r < R$, $\varepsilon = R - r > 0$. Then $y \in B(x, \varepsilon)$ $\Rightarrow \|y - x\| < \varepsilon \Rightarrow \|y - a\| < \|y - x\| + \|x - a\| < \varepsilon + r = R$ So $B(x, \varepsilon) \subset B(a, R)$.

Consider $\mathbb{R}$. Then E is open in $\mathbb{R} \Leftrightarrow \forall x \in E$, $\exists \varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subset E$. So E is a (possibly infinite) union of open intervals. Every such union is open.

$E = \mathbb{Q} \subset \mathbb{R}$ is not open, since $\forall \varepsilon > 0$, $(-\varepsilon, \varepsilon)$ contains an irrational, but is not closed since its compliment is not open, because $(\sqrt{2} - \varepsilon, \sqrt{2} + \varepsilon)$ always contains a rational.

Definition

Let $(V, \|\cdot\|)$, $(V', \|\cdot\|')$ be normed spaces, and $E \subset V$ a subset, $f: E \to V$ a mapping. We say that f is continuous at $x \in E$ if:

$\forall \varepsilon > 0$, $\exists \delta > 0$ such that $y \in E$, $\|x-y\| < \delta \Rightarrow \|f(x) - f(y)\|' < \varepsilon$
f is continuous if it is continuous at each $x \in E$. If $V' = \mathbb{R}$, we say that f is a continuous function on E.

Equivalently, f is continuous at x if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $f^{-1}(B(f(x), \varepsilon)) \supset E \cap B(x, \delta)$
$f^{-1}(Y) = \{x \in E \mid f(x) \in Y\}$

If $E = V$, this means that the pre-image of any open set in V' is an open set in V.

Theorem 15

Let $E \subset V$, $f: E \to V'$ as above. Then f is continuous at $x \in E$ $\Leftrightarrow$ for every sequence $(x_n) \in E$ with $(x_n) \to x$, $(f(x_n)) \to f(x)$ in V'.

Proof

i) Assume that f is continuous at $x \in E$. Let $x_n \to x$. Let $\varepsilon > 0$. Then $\exists \delta > 0$ such that $\|y - x\| < \delta$, $y \in E \Rightarrow \|f(y) - f(x)\|' < \varepsilon$. Also, $\exists N_\delta$ such that $\forall n \geq N_\delta$, $\|x - x_n\| < \delta$, and so $\|f(x_n) - f(x)\|' < \varepsilon$ i.e. $f(x_n) \to f(x)$

ii) Assume that f is not continuous at $x \in E$. So $\exists \varepsilon > 0$ such that $\forall \delta > 0$, $\exists y \in E$ with $\|y - x\| < \delta$, but $\|f(y) - f(x)\|' \geq \varepsilon$. Let $\delta = 1/n$ $(n \geq 1)$. $\exists x_n \in E$ with $\|x_n - x\| < 1/n$ but $\|f(x_n) - f(x)\| \geq \varepsilon$.
Then $(x_n) \to x$, but $(f(x_n)) \not\to f(x)$.

A simple but important example of a continuous function is:

Proposition 16

Let $(V, \|\cdot\|)$ be a normed space. Then $f: V \to \mathbb{R}$, $f(x) = \|x\|$ is continuous.

Proof

Note that $\forall x, y \in V$, $\|x - y\| \geq \big|\|x\| - \|y\|\big|$
Let $x \in V$, $\varepsilon > 0$, $f: V \to \mathbb{R}$, $f(x) = \|x\|$.
Then $\forall y \in V$:

$$\|y - x\| < \varepsilon \Rightarrow \underbrace{\big|\|y\| - \|x\|\big|}_{= |f(x) - f(y)|} < \varepsilon \qquad (\delta = \varepsilon)$$

So f is continuous at x.

Remarks

1. If $E \subset V$, $f: E \to V'$ is continuous, and $E' \subset E$ let $g: E' \to V'$ be the mapping $g(x) = f(x)$ $\forall x \in E'$. Here, g is the restriction of f to E'. It follows from the definition that g is also continuous.

2. $f, g : E \to V'$, continuous, $a, b \in \mathbb{R}$. Then $af + bg$ is also continuous.

3. V, V', V'' normed spaces. $E \subset V$, $E' \subset V'$.
$f : E \to V'$ with $f(E) \subset E'$, $g : E' \to V''$.
Then f, g continuous $\Rightarrow$ $g \circ f$ is continuous.

4. $f : E \to \mathbb{R}^n$ is continuous $\Leftrightarrow$ Each component $f_i : E \to \mathbb{R}$ is continuous
(Theorem 15 combined with proposition 7)

Analysis II ⑩

Definition

Let V, V' be normed spaces, $E \subset V$. We say that $f : E \to V'$ is <u>uniformly continuous</u> (on E) if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $x, y \in E$, $\|x - y\| < \delta \Rightarrow \|f(x) - f(y)\|' < \varepsilon$
If $V = V' = \mathbb{R}$, we recover our earlier definition.

The following is an important class of uniformly continuous functions

Definition

$V \supset E \to V'$ as above. We say that f is <u>Lipschitz</u> if $\exists$ $c > 0$ such that $\forall x, y \in E$, $\|f(x) - f(y)\|' \leq c \|x - y\|$
Note that if f is Lipschitz, then it is uniformly continuous (given $\varepsilon > 0$, take any δ with $0 < \delta < \frac{\varepsilon}{c}$).

Example

$f : [0,1] \to \mathbb{R}$, $f(x) = \sqrt{x}$. Then f is continuous, so it is uniformly continuous (since $[0,1]$ is closed and bounded). But f is not Lipschitz.

$$\frac{\|f(x) - f(0)\|}{\|x - 0\|} = \frac{x^{\frac{1}{2}}}{x} = x^{-\frac{1}{2}}$$ which is <u>unbounded</u> for $x \in [0,1]$
hence no such c exists.

$x \neq y$, f Lipschitz $\Rightarrow \frac{\|f(x) - f(y)\|}{\|x - y\|} < c$, for $V = V' = \mathbb{R}$.
Compare this with differentiability:

$$\frac{f(y) - f(x)}{y - x} \to f'(x) \text{ as } y \to x$$

Theorem 17

Let $E \subset \mathbb{R}^m$ be a closed, bounded subset, and $f: E \to \mathbb{R}^n$ be continuous. Then

i) f is uniformly continuous.

ii) $f(E)$ is closed and bounded.

Lemma

Let $E \subset \mathbb{R}^n$ be any subset. Then E is closed and bounded $\Leftrightarrow$ every sequence in E has a subsequence which converges to an element of E.

Proof

($\Rightarrow$) Bolzano-Weierstrass says that any sequence in E has a subsequence converging to an element of $\mathbb{R}^n$. As E is closed, the limit must also be in E.

($\Leftarrow$) Assume that every sequence in $\mathbb{R}$ has such a subsequence. Then, E is bounded, because if not, $\exists\, (x^{(k)})$ in E with $\|x^{(k)}\| \to \infty$, but this cannot have a convergent subsequence. Also, if $(x^{(k)}) \to x$, where $x^{(k)} \in E$, $x \in \mathbb{R}^n$, then any subsequence of $(x^{(k)})$ converges to x, so by assumption, $x \in E$. This shows that every limit point of E belongs to E, so E is closed.

In other words, for $E \subset \mathbb{R}^n$, E is closed and bounded $\Leftrightarrow$ E is sequentially compact.

Analysis II ⑩

Proof of Theorem 17

i) Take the proof for $f : [a,b] \to \mathbb{R}$ and replace $|\cdot|$ by $\|\cdot\|$ everywhere.

ii) Let $(y^{(k)})$ be any sequence in $f(E)$. It is sufficient to prove that $(y^{(k)})$ has a subsequence converging to an element of $f(E)$. $y^{(k)} = f(x^{(k)})$ for some $x^{(k)} \in E$. As E is closed and bounded, $\exists (x^{\sigma(k)}) \to x \in E$. But f is continuous, so $(y^{\sigma(k)}) = (f(x^{\sigma(k)})) \to f(x) \in f(E)$.

Corollary

$E \subset \mathbb{R}^m$ closed and bounded, $f : E \to \mathbb{R}$ continuous. Then, for $E \neq \emptyset$, f is uniformly continuous, f is bounded on E, and attains its bounds.

Proof

The only non-trivial part is that "f attains its bounds". But by ii), f is closed (and bounded). Then if $b = \sup f(E)$, either $b \in f(E)$ or b is a limit point of $f(E)$ by the definition of sup, so as $f(E)$ is closed, $b \in f(E)$ either way. So $\sup f(E) \in f(E)$ and likewise $\inf f(E) \in f(E)$, i.e. f attains its bounds.

Important Remarks

1. It is essential to be in $\mathbb{R}^n$ here. This does not hold for closed and bounded subsets of an arbitrary normed space.

2. If $E \subset \mathbb{R}^n$ closed and bounded, then $f(E)$ is also closed and bounded.

But, if E is merely closed, $f(E)$ needn't be closed e.g.
$E = \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{1+x^2}$
E is closed, but $f(E)$ is $(0, 1]$, not closed.

If E is merely bounded, $f(E)$ needn't be bounded e.g.
$E = (0,1) \to \mathbb{R}$, $f(x) = \frac{1}{x}$, $f(E) = (1, \infty)$, unbounded

The key to this is <u>compactness</u>. A subset $E \subset \mathbb{R}^n$ is <u>compact</u> $\Leftrightarrow$ it is closed and bounded, and in any topological space, continuous functions preserve compactness.

Application:

<u>Theorem 18</u>

Any two norms on $\mathbb{R}^n$ are equivalent.

<u>Proof</u>

Let $\|\cdot\|$ be Euclidean norm on $\mathbb{R}^n$, and let $f: \mathbb{R}^n \to \mathbb{R}$ be any norm. It is sufficient to prove that $\|\cdot\|$ and f are equivalent.

Let $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ be the i^{th} basis vector in $\mathbb{R}^n$, and $c_i = f(e_i) > 0$. Let $x = \sum x_i e_i \in \mathbb{R}^n$ and choose $y = (\pm c_1, \dots, \pm c_n)$ with signs such that $y_i x_i = c_i |x_i|$

Then $f(x) = \sum_k f(x_k e_k) = \sum_k c_k |x_k| = x \cdot y$ (as f is a norm
$\leq \|x\| \|y\| = B\|x\|$, $B = \|y\|$

So if $x, y \in \mathbb{R}^n$, $\|x-y\| \geq \frac{1}{B} f(x-y)$
$\geq \frac{1}{B} |f(x) - f(y)|$
by the triangle inequality for f.

$\therefore \|x-y\| < \delta \Rightarrow |f(x)-f(y)| < \varepsilon/B$

So we have shown that $f: \mathbb{R}^n$ is <u>continuous</u> on $(\mathbb{R}^n, \|\cdot\|_{\text{Euclid}}$

02/11/11

Analysis II (11)

Last time

$\|\cdot\|$, f norms on $\mathbb{R}^n$ ($\|\cdot\|$ Euclidean norm).

We showed that if $c_i = f(0, \dots, 0, 1, 0, \dots, 0)$

then $\forall x \in \mathbb{R}^n$ $f(x) \leq B\|x\|$, $B = (\sum c_i^2)^{\frac{1}{2}}$

$\Rightarrow \forall x, y \in \mathbb{R}^n$, $\|x-y\| < \delta \Rightarrow |f(x) - f(y)| \leq B\delta$

So $f: \mathbb{R}^n \to \mathbb{R}$ is continuous (for Euclidean norm) - actually is a Lipschitz mapping

Let $E = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ which is bounded (obvious) and closed (easy to see). Now $f: E \to \mathbb{R}$ is continuous, and $f(x) \neq 0$ $\forall x \in E$ as f is a norm. So f is bounded and attains its bounds on E. Let $A = \inf |f(E)|$ then $\exists x \in E$ such that $f(x) = A$ (remember $f \geq 0$)

So $A > 0$, hence $\forall x \in \mathbb{R}^n$, $f(x) \geq A\|x\|$

(if $\|x\| = 1$, $x \in E$; if $\|x\| \neq 1$, ~~$x \neq 0$, replace~~ then $x = \|x\| y$ for some $y \in E$, then $f(y) \geq A \Rightarrow f(x) \geq A\|x\|$

$\exists A, B > 0$ such that $A\|x\| \leq f(x) \leq B\|x\|$ $\forall x \in \mathbb{R}^n$

i.e. $\|\cdot\|$, f are equivalent.

Norms on $L(\mathbb{R}^m, \mathbb{R}^n)$

Let $A: \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then A is continuous.

It is sufficient to check for $A_i: \mathbb{R}^m \to \mathbb{R}$, $A_i x = \sum_{j=1} a_{ij} x_j$, but then it is enough to check for the functions $(x_1, \dots, x_n) \mapsto x_j$ which is then easy to see.

Consider $(\mathbb{R}^m, \mathbb{R}^n) = \{$linear transformations $\mathbb{R}^m \to \mathbb{R}^n\} \cong \mathbb{R}^{mn}$ as a vector space (think $m \times n$ matrices)

Let $E = \{x \in \mathbb{R}^m \mid \|x\| = 1\}$

So $A(E)$ is bounded as A is continuous, E closed and bounded.

Defined <u>operator norm</u> on $L(\mathbb{R}^m, \mathbb{R}^n)$ by $\|A\| = \sup_E \|Ax\|$

$\|A\| = \sup_{0 \neq x \in \mathbb{R}^n} \left(\frac{\|Ax\|}{\|x\|}\right)$ (as in proof of Theorem 18)

(1) This is a norm: only non-obvious property is the triangle inequality

$A, B \in L(\mathbb{R}^m, \mathbb{R}^n)$

$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\| + \|B\|$ if $x \in E$

$\Rightarrow \|A+B\| \leq \|A\| + \|B\|$

2. $\forall x \in \mathbb{R}^m$, $\|Ax\| \leq \|A\| \|x\|$ (clear from the definition)
3. $A : \mathbb{R}^m \to \mathbb{R}^n$, $B : \mathbb{R}^n \to \mathbb{R}^p$ $\Rightarrow \|BA\| \leq \|B\| \|A\|$ (easy exercise)

2 is a special case of 3 since if $m=1$, $A : \mathbb{R} \to \mathbb{R}^n$, $\lambda \mapsto \lambda v$ where $v = A(1)$, then $\|A\|_{\text{operator norm}} = \|v\|_{\text{Euclidean}}$

Any two norms on this space are equivalent by Theorem 18; in particular "Euclidean norm on matrices" $A = (a_{ij})$

$$\|A\| = \left(\sum_{ij} (a_{ij})^2\right)^{\frac{1}{2}}$$

Operator norm has the advantage of property 3.

3 Differentiation in Euclidean Space

Digression on limits in $\mathbb{R}^n$

<u>Definition</u> Let $E \subset \mathbb{R}^m$, $a \in \mathbb{R}^m$, such that for some $r > 0$, $\{x \in \mathbb{R}^m \mid 0 < \|x-a\| < r\} \subset E$. Let $f : E \to \mathbb{R}^n$ be a mapping and $b \in \mathbb{R}^n$. We say $f(x) \to b$ as $x \to a$ or $b = \lim_{x\to a} f(x)$ if $\forall \varepsilon > 0$, $\exists \delta$ with $0 < \delta \leq r$ such that

$$0 < \|x-a\| < \delta \Rightarrow \|f(x) - b\| < \varepsilon$$

(so $f(x)$ is defined) We need f to be defined at all points sufficiently close to a.

Limits enjoy similar properties to limits in $\mathbb{R}$

i.e. $\lim (f+g) = \lim f + \lim g$

$f : E \to \mathbb{R}^n$, $g : E \to \mathbb{R}$, $\lim fg = (\lim f)(\lim g)$

$\lim f = (\lim f_1, \lim f_2, \ldots, \lim f_n)$ (same as for sequences.

We say that f is continuous at $x = a$ if $a \in E$ and $\lim_{x\to a} f(x) = f(a)$

Analysis II (11)

$f : (b,c) \to \mathbb{R}$. Recall that f is differentiable at $a \in (b,c)$ if $f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$ exists; equivalently, f is differentiable at $x = a$ iff:
$\exists A \ (= f'(a))$ such that $\frac{f(a+h) - f(a) - Ah}{h} \to 0$ as $h \to 0$
So $f(a) + hA$ is the "best linear approximation" to $f(a+h)$ as $h \to 0$.

i.e. $f(x) - [f(a) + A(x-a)] \to 0$ "more rapidly than $x \to a$"

For a function $\mathbb{R}^m \to \mathbb{R}^n$, we will define the derivative to be the "best possible linear approximation".

<u>Definition</u> Let $U \subset \mathbb{R}^m$ be an open subset, $f : U \to \mathbb{R}^n$, $a \in U$.
We say that <u>f is differentiable at a</u> if

$\exists A \in L(\mathbb{R}^m, \mathbb{R}^n)$ such that $\frac{f(a+h) - f(a) - Ah}{\|h\|} \to 0$

as $h \to 0$ ($h \in \mathbb{R}^m$).

We say that A is the <u>derivative</u> of f at A.

(possible notations, $f'(a)$, $(Df)(a)$, $D_a f$, Df_a)

We require U to be open to ensure that $\exists r > 0$ with $B(a,r) \subset U$.

04/11/11

Analysis II (12)

Last time $\mathbb{R}^m \underset{\text{open}}{\supseteq} U \xrightarrow{f} \mathbb{R}^n$

Then f is differentiable at $a \in U$ if $\exists A : \mathbb{R}^m \to \mathbb{R}^n$, a linear transformation, such that ($h \in \mathbb{R}^m$ restricted $a+h \in U$)

$$\frac{1}{\|h\|}\left[f(a+h) - f(a) - Ah\right] \to 0 \text{ as } h \to 0$$

Proposition 19

i) If f is differentiable at $a \in U$ then $(Df)(a)$ is unique.

ii) If f is differentiable at a, then f is continuous at a.

iii) $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$, $f_i : U \to \mathbb{R}$. f is differentiable at a $\Leftrightarrow$ all the f_i are differentiable at a, and:

$$Df(a) = \begin{pmatrix} Df_1(a) \\ \vdots \\ Df_n(a) \end{pmatrix} \in L(\mathbb{R}^m, \mathbb{R}^n), \quad Df_i(a) : \mathbb{R}^m \to \mathbb{R}$$

Proof

i) Suppose A, A' are both derivatives of f at a. Fix $h \in \mathbb{R}^m$, $\neq 0$, then

$$\frac{f(a+\lambda h) - f(a) - A \cdot \lambda h}{\lambda \|h\|} \to 0 \text{ as } \lambda \to 0$$

i.e. $$\frac{f(a+\lambda h) - f(a)}{\lambda} \to Ah \text{ as } \lambda \to 0 \quad (*)$$

But also $\dfrac{f(a+\lambda h) - f(a)}{\lambda} \to A'h$ as $\lambda \to 0$, so $\forall h$, $Ah = A'h$ i.e. $A = A'$

ii) $f(a+h) - f(a) - Df(a)h \to 0$ as $h \to 0$

$\Rightarrow f(a+h) \to f(a)$ as $h \to 0$ i.e. f continuous at A

iii) If $A_i = Df_i(a) \in L(\mathbb{R}^m, \mathbb{R})$ exists $\forall i$ ← row vectors

and $A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \in L(\mathbb{R}^m, \mathbb{R}^n)$

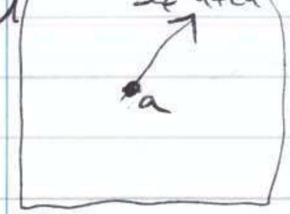
then $$\frac{f(a+h) - f(a) - Ah}{\|h\|} = \left(\frac{f_i(a+h) - f_i(a) - A_i h}{\|h\|}\right)_i \to 0$$ (vector)

since each component $\to 0$

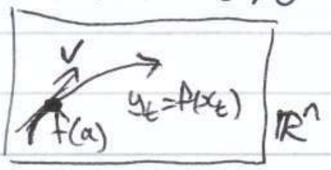
So by iii) it is often enough to consider $f : U \to \mathbb{R}$ ($U \subset$ open $\mathbb{R}^m$)

Directional derivatives $\mathbb{R}^m \supset U \xrightarrow{f} \mathbb{R}^n$ ($n=1$ if you like)

Let $u \in \mathbb{R}^m$, directional derivative along u at a is (if it exists):



$$v = \frac{d}{dt} f(x_t)\Big|_{t=0} = \lim_{t \to 0} \frac{f(a+tu) - f(a)}{t}$$



(see (*) in proof ii))

<u>Relation 6</u>: If f is differentiable at $a \in U$, then the directional derivative along u is just $\underbrace{Df(a)}_{L(\mathbb{R}^m, \mathbb{R}^n)} \underbrace{(u)}_{\mathbb{R}^m} \in \mathbb{R}^n$

(If $m=1$ and $u = 1 \in \mathbb{R}$, then the directional derivative along u is just the usual derivative. The proof is in equation (*) by replacing λ, h by t, u).

<u>In particular</u>, as $Df(a)$ is <u>linear</u>, if f is differentiable, its directional derivative depends <u>linearly</u> on u (this is a <u>strong</u> property).

<u>Partial derivatives</u> are just directional derivatives along the coordinate axes.

$$\frac{\partial f}{\partial x_j}(a) = \frac{d}{dt} f(a_1, \dots, a_j + t, \dots, a_m)\Big|_{t=0} = \text{directional derivative along } \underline{e_j}$$

<u>Proposition 20</u> $\mathbb{R}^m \supset U$ open, $f: U \to \mathbb{R}^n$ differentiable at $a \in U$. Then the partial derivative $\frac{\partial f_i}{\partial x_j}$ at $x=a$ exist, and the matrix of $Df(a)$ is just $\left(\frac{\partial f_i}{\partial x_j}\right)_{ij}$

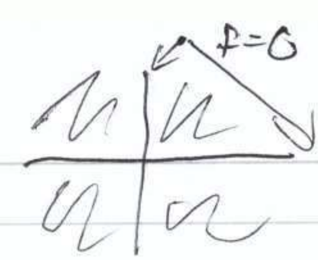
<u>Proof</u> ($n=1$) We have seen that $\frac{\partial f}{\partial x_j}(a)$ is the directional derivative of f along $\underline{e_j}$, so $\frac{\partial f}{\partial x_j}(a) = \underbrace{(Df(a))}_{L(\mathbb{R}^m, \mathbb{R})} \underline{e_j} \in \mathbb{R}$

(as a row vector) $n > 1$ dealt with by proposition 19 iii)

2

04/11/11

Analysis II (12)



Example

$f:\mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \begin{cases} 0 & \text{if } xy = 0 \\ 1 & xy \neq 0 \end{cases}$

At $(0,0)$, f is not continuous, so certainly not differentiable. W f=

But $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ exist at $(0,0)$ and equal 0

$$\frac{\partial f}{\partial x}(0,0) = \frac{d}{dt} f(t,0)\Big|_{t=0} = 0$$

($\frac{\partial f}{\partial x}(0)$ only "knows" about f along the x axis)

Note that directional derivatives in other directions do not exist.

See sheet 3 for a function which has all directional derivatives at some point a but is not differentiable at a.

Easy properties of the derivative:

Proposition 21

i) Let $f:\mathbb{R}^m \to \mathbb{R}^n$, $f(x) = Ax$ for $A \in L(\mathbb{R}^m, \mathbb{R}^n)$
Then f is differentiable $\forall a \in \mathbb{R}^m$, $Df(a) = A$

ii) $\mathbb{R}^m \supset U \xrightarrow{f} \mathbb{R}^n$ differentiable at $a \in U$
$\mathbb{R}^p \supset V \xrightarrow{g} \mathbb{R}^q$ " " $b \in V$
Then $\mathbb{R}^m \times \mathbb{R}^p \supset U \times V$ is open and $(f,g): U \times V \to \mathbb{R}^{n+q}$
$(u,v) \mapsto (f(u), g($

is differentiable at $(a,b) \in (U \times V)$ with derivative

$$(Df)(a) \oplus (Dg)(b) \quad \text{or} \quad \left(\begin{array}{c|c} Df(a) & 0 \\ \hline 0 & Dg(b) \end{array}\right)$$

Exercise in the definition.

3

Analysis II (13)

<u>Theorem 22 ("Chain rule")</u> $\mathbb{R}^m \underset{\text{open}}{\supset} U \xrightarrow{f} \mathbb{R}^n$

$\mathbb{R}^n \underset{\text{open}}{\supset} V \xrightarrow{g} \mathbb{R}^p$, Let $a \in U$ such that $f(a) = b \in V$, assume f is differentiable at a and g at b. Then $g \circ f$ is defined in some ball containing a, and is differentiable there with derivative :

$$D(g \circ f)(a) = Dg(b) \circ Df(a) : \mathbb{R}^m \to \mathbb{R}^p$$

<u>Proof</u>

f is continuous at $a \Rightarrow \forall \varepsilon > 0, \exists \delta > 0$ such that $\forall x \in U$ with $\|x - a\| < \delta$, $\|f(x) - f(a)\| < \varepsilon \Rightarrow$ if we choose $\varepsilon > 0$ such that $B(b, \varepsilon) \subset V$, then $\forall x \in B(a, \delta)$ (which WLOG is contained in U), $f(x) \in B(b, \varepsilon) \subset V$, so $g \circ f$ is defined on $B(a, \delta)$.

Write $A = Df(a) \in L(\mathbb{R}^m, \mathbb{R}^n)$; $B = Dg(b) \in L(\mathbb{R}^n, \mathbb{R}^p)$

$$k(h) = \begin{cases} \frac{1}{\|h\|}[f(a+h) - f(a) - Ah] & \text{if } 0 \neq h \in \mathbb{R}^m \text{ with } a+h \in U \\ 0 & \text{if } h = 0 \in \mathbb{R}^m \end{cases}$$

$$s(k) = \begin{cases} \frac{1}{\|k\|}[g(b+k) - g(b) - Bk] & \text{if } 0 \neq k \in \mathbb{R}^n \text{ with } b+k \in V \\ 0 & \text{if } k = 0 \in \mathbb{R}^n \end{cases}$$

So by differentiability of f and g, s and k are continuous at 0.

$$(g \circ f)(a+h) = g(f(a) + Ah + \|h\| r(h))$$

$$= g(f(a)) + \underbrace{B(Ah + \|h\| r(h))}_{\cancel{=Bk}} \cancel{+ \|k\| s(k)} + \|Ah + \|h\| r(h)\| \times s(Ah + \|h\| r(h))$$

(*)

$$= g(f(a)) + Bk + \|k\| s(k)$$

$$\frac{1}{\|h\|}[(g \circ f)(a+h) - (g \circ f)(a) - BAh] = Br(h) + \left\|\frac{Ah}{\|h\|} + r(h)\right\| s(Ah + \|h\| r(h))$$

Now let $h \to 0$.

$r(h) \to 0$ and $s(Ah + \|h\| r(h)) \to 0$.

and $\left\|\frac{Ah}{\|h\|}\right\| \leq$ operator norm of A, so is bounded as $h \to 0$

So (*) $\to 0$ as $h \to 0$, and hence $g \circ f$ is differentiable at a with derivative BA.

(I.e. The derivative of composite functions is the composite of the derivatives at the appropriate points)

Alternatively, write $o(x)$ for any function for which $\frac{o(x)}{\|x\|} \to 0$ as $x \to 0$.
e.g. $(Df)(a)$ is characterised by $f(a+h) = f(a) + Df(a)h + o(h)$
and $(Dg)(b)$ by $g(b+k) = g(b) + Dg(b)k + o(k)$
So $(g \circ f)(a+h) = g(b + Ah + o(h)) = g(b) + B(Ah + o(h)) + \underbrace{o(Ah + o(h))}_{o(h)}$

$$= g(b) + BAh + o(h)$$

i.e. $D(g \circ f)(a) = BA$

<u>Corollary</u> $f, g: U \to \mathbb{R}^n$ differentiable at $a \in U \subset_{\text{open}} \mathbb{R}^n$
$\lambda, \mu \in \mathbb{R} \Rightarrow \lambda f + \mu g$ is differentiable at a, where
$D(\lambda f + \mu g)(a) = \lambda Df(a) + \mu Dg(a)$

<u>Proof</u>
$\lambda f + \mu g$ is the composite of 3 maps:

$$\mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{2m} \supset U \times U \xrightarrow{f \oplus g} \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

$$x \longmapsto (x, x) \qquad (x, y) \longmapsto (f(x), g(x)) \qquad (u, v) \mapsto \lambda u + \mu v$$

"diagonal map"

The outer two maps are linear, and the middle map is differentiable at (a, a) with derivative $(Df)(a) \oplus (Dg)(b)$ (Prop 2, part ii))
so apply the chain rule. ~~$D(\lambda f + \mu g)(a) =$~~

$$D(\lambda f + \mu g)(a) = (\lambda I_n + \mu I_n) \begin{pmatrix} Df(a) & 0 \\ 0 & Dg(a) \end{pmatrix} \begin{pmatrix} I_m \\ I_m \end{pmatrix} = \lambda Df(a) + \mu Dg(a)$$

<u>Theorem 23</u>
Let $U \subset \mathbb{R}^n$ be open, $f: U \to \mathbb{R}$, $a \in U$.
Suppose that:
i) For some $B(a, r) \subset U$, the partial derivatives $\frac{\partial f}{\partial x_j}$ $(1 \le j \le m)$ exist at all points in $B(a, r)$
ii) $\left\{\frac{\partial f}{\partial x_j}\right\}$ are continuous at $x = a$.
Then, f is differentiable at $x = a$, with $Df(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)$

Proof $\quad \rightarrow a = (a_1, a_2) \in U \subset \mathbb{R}^2$

First consider $m = 2$. By the Mean Value Theorem,

$$f(a_1 + h_1, a_2 + h_2) = f(a_1, a_2 + h_2) + h_1 \frac{\partial f}{\partial x_1}(b_1, a_2 + h_2)$$

(for some b_1 with $|b_1 - a_1| < |h_1|$)

$$= f(a_1, a_2) + h_2 \frac{\partial f}{\partial x_2}(a_1, b_2) + h_1 \frac{\partial f}{\partial x_1}(b_1, a_2 + h_2)$$

(for some b_2 with $|b_2 - a_2| < |h_2|$)

So $\frac{1}{\|h\|} \left| f(a+h) - f(a) - \left(\frac{\partial f}{\partial x_1}(a) h_1 + \frac{\partial f}{\partial x_2}(a) h_2 \right) \right|$

$$(*) \quad = \frac{1}{\|h\|} \left| h_1 \left(\frac{\partial f}{\partial x_1}(b_1, a_2 + h_2) - \frac{\partial f}{\partial x_1}(a) \right) + h_2 \left(\frac{\partial f}{\partial x_2}(a_1, b_2) - \frac{\partial f}{\partial x_2}(a) \right) \right|$$

$$\leq \left| \frac{\partial f}{\partial x_1}(b_1, a_2 + h_2) - \frac{\partial f}{\partial x_1} \right| + \left| \frac{\partial f}{\partial x_2}(a_1, b_2) - \frac{\partial f}{\partial x_2}(a) \right| \quad \text{as } \|h\| \geq |h_i|, \ i = 1, 2$$

As both partial derivatives at continuous at $x = a$, given $\varepsilon > 0$, $\exists \delta$ such that if $\left| \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a) \right| < \frac{\varepsilon}{2}$

So provided $\|h\| < \delta$, $(*) < \varepsilon$, so f is differentiable at $x = a$.

Say $f : U \to \mathbb{R}^n$, $U \underset{\text{open}}{\subset} \mathbb{R}^m$, is <u>continuously differentiable</u> if Df exists everywhere on U and is <u>continuous</u> (as a function $Df : U \to L(\mathbb{R}^m, \mathbb{R}^n)$)

<u>Corollary</u> Let $f = (f_1, \dots, f_n) : U \to \mathbb{R}^n$ ~~be continuously~~

Then f is continuously differentiable on U iff all partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist everywhere on U and are continuous.

(Proof: for $n = 1$ this is Theorem 23, in general ~~apply Proposition~~ reduce to individual coordinates).

Analysis II (14)

Proof of Theorem 23

(continuous partial derivatives $\Rightarrow$ differentiability)

$\mathbb{R}^n \overset{\text{open}}{\supset} U \xrightarrow{f} \mathbb{R}$, $U \ni a$, $\frac{\partial f}{\partial x}$ exist and are continuous at $x = a$

General case: induction on m. We may assume that the function $f(x_1, \ldots, x_{m-1}, a_m)$ is differentiable at $(x_1, \ldots, x_{m-1}) = (a_1, \ldots, a_{m-1})$ with derivative $\left(\frac{\partial f}{\partial x_j}(a)\right)_{1 \le j \le m-1}$. So if $h = (h_1, \ldots, h_m)$, $\|h\| < r$ then $f(a_1 + h_1, \ldots, a_{m-1} + h_{m-1}, a_m)$

$$= f(a) + \sum_{j=1}^{m-1} \left(\frac{\partial f}{\partial x_j}(a) h_j \right) + o(h)$$

Also $f(a_1 + h_1, \ldots, a_m + h_m) = f(a_1 + h_1, \ldots, a_{m-1} + h_{m-1}, a_m)$
$+ \frac{\partial f}{\partial x_m}(a_1 + h_1, \ldots, a_{m-1} + h_{m-1}, a_m)$
$+ o(h_m)$

and $\frac{\partial f}{\partial x_m}(a_1 + h_1, \ldots, a_{m-1} + h_{m-1}, a_m) - \frac{\partial f}{\partial x_m}(a) \to 0$ as $h \to 0$, by assumption

so (" " " " ") $h_m = o(h)$

Therefore $f(a+h) = f(a) + \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(a) h_j + o(h)$ as required.

Remark $m = 1$, n arbitrary

$f : (a, b) \to \mathbb{R}^n$ differentiable at $x \in (a, b)$

$f = (f_1, \ldots, f_n)$, then $Df(x) \in L(\mathbb{R}, \mathbb{R}^n) \cong \mathbb{R}^n$ (canonical isomorphism, $(L : \mathbb{R} \to \mathbb{R}^n) \mapsto L(1)$)

i.e. $Df : (a, b) \to \mathbb{R}^n$, $Df = \begin{pmatrix} f_1' \\ \vdots \\ f_n' \end{pmatrix}$, f_i' "usual" derivative

$x \mapsto f(x)$ is a parametrised curve in $\mathbb{R}^n$

$Df(x)$ (if non-zero) is a tangent vector to this curve $f(x)$.

Mean value Inequalities For $f : [a, b] \to \mathbb{R}$, continuous, differentiable on (a, b) we have the Mean Value Theorem:

$f(b) - f(a) = f'(\xi)(b - a)$ for some $\xi \in (a, b)$

Simplest generalisation:

Theorem 24 $f : [a, b] \to \mathbb{R}^n$ continuous, differentiable on (a, b)

Suppose $\|Df(x)\| \le k$ $\forall x \in (a, b)$.

Then $\|f(b) - f(a)\| \le k(b - a)$

Proof of Theorem 24

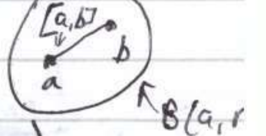
$F = (f_1, \dots, f_n)$, $f_i : [a,b] \to \mathbb{R}$. Let $w = f(b) - f(a) \in \mathbb{R}^n$
and $g : [a,b] \to \mathbb{R}$ be the function $g(x) = \langle w, f(x) \rangle = \sum w_i f_i$
Chain rule $\Rightarrow$ g is differentiable on (a,b) with
$g'(x) = \sum w_i \frac{\partial f_i}{\partial x} = \langle w, Df \rangle$. By the Mean Value Theorem on g
$\exists\, \xi \in (a,b)$ such that $g(b) - g(a) = (b-a)\, g'(\xi)$

$$g(b) - g(a) = \langle w, f(b) \rangle - \langle w, f(a) \rangle = \|f(b) - f(a)\|^2 = \|w\|^2$$

$$|(b-a)\, g'(\xi)| = (b-a)\,|\langle w, Df(\xi) \rangle| \le (b-a)\,\|w\|\,\|Df(\xi)\| \le (b-a)\, k \|w\|$$

Either $\|w\| = 0$, or $\|w\| = \|f(b) - f(a)\| \le (b-a)k$
(Compare with the proof that $\| \int f \| \le \int \|f\|$)

Theorem 25 $\mathbb{R}^n \supset B(a,r) \xrightarrow{f} \mathbb{R}^n$, differentiable. Let $b \in B(a,r)$ and assume that the derivative $\|Df\| \le k$ (operator norm) on $[a,b] = \{ ta + (1-t)b \mid t \in [0,1] \} \subset B(a,r)$
Then $\|f(b) - f(a)\| \le \|b-a\|\, k$



(in applications, typically, we will have the stronger statement $\|Df\| \le k$ on $B(a,r)$, not needed for the proof)

Proof
Consider $g : [0,1] \to \mathbb{R}^n$, $g(x) = f[(1-x)a + xb]$
which is the composite $f \circ r$, $r : [0,1] \to \underset{\cup B(a,r)}{\mathbb{R}^n}$, $r(x) = (1-x)a + xb$

Chain rule $\Rightarrow$ g is differentiable on $(0,1)$. $Dg(x) = Df(r(x))\, Dr(x)$
$$= \underbrace{Df(r(x))}_{L(\mathbb{R}^n, \mathbb{R}^n)} \underbrace{(b-a)}_{\in \mathbb{R}^n}$$

~~So by Theorem 24 $\|Dg(x)\| = \|$~~
So $\|Dg(x)\| = \|Df(r(x))(b-a)\| \le \|Df(r(x))\| \cdot \|b-a\|$
(by definition of the operator norm), so by Theorem 24 (applied to g)

$$\|f(b) - f(a)\| = \|g(1) - g(0)\| \le k \|b-a\|$$

Analysis II (14)

Corollary

Let $f: B(a,r) \to \mathbb{R}^n$ $(a \in \mathbb{R}^m)$ be differentiable on $B(a,r)$ with $Df = 0$ on $B(a,r)$. Then f is constant on $B(a,r)$ i.e. $f(x) = f(a)$ $\forall x \in B(a,r)$.

Proof

Apply Theorem 25 to f. We may take $k = 0$. So $\forall x \in B(a,r)$
$\|f(x) - f(a)\| \leq 0 \Rightarrow f(x) = f(a)$.

We would like to replace $B(a,r)$ by any open set $U \subset \mathbb{R}^m$. But the obvious generalisation fails ~~because~~ (even if $m = 1$).

E.g. $U = (0,1) \cup (2,3) \subset \mathbb{R}$

$f: U \to \mathbb{R}$, $f(x) = \begin{cases} 0 & \text{if } x \in (0,1) \\ 1 & \text{if } x \in (2,3) \end{cases}$

Then f is differentiable on U with $Df = 0$, but f is not constant. (it is however locally constant). $\to$ need path connectedness

11/11/11

Analysis II (15)

$f: B(a,r) \to \mathbb{R}^n$ differentiable, $Df = 0$ on $B(a,r) \Rightarrow f$ constant

Definition

Let $E \subset \mathbb{R}^m$ be a non-empty subset. We say that E is path-connected if $\forall a, b \in E$, $\exists$ a continuous $\gamma : [0,1] \to E$ such that $\gamma(0) = a$, $\gamma(1) = b$.

(intuitively, a and b can be joined by a continuous curve contained in E)

Proposition 2.6 The path-connected subsets of $\mathbb{R}$ are precisely the intervals (a,b), $(a,b]$, $[a,b)$, $[a,b]$ $(-\infty \le a \le b \le \infty)$

Proof

Clearly, any interval is path connected. Conversely, let E be path connected, and let $a = \inf E$, $b = \sup E$ ($\pm\infty$ are allowed). We claim that $a < x < b \Rightarrow x \in E$, and if so then E is one of the intervals above. If $x > a$, $\exists a_1 \in E$ with $a_1 < x$. If $x < b$, $\exists b_1 \in E$ with $x < b_1$. As E is path connected, $\exists \gamma : [0,1] \to E$, continuous, with $\gamma(0) = a_1$, $\gamma(1) = b_1$. By the Intermediate Value Theorem, since $a_1 < x < b_1$, $\exists t \in (0,1)$ with $\gamma(t) = x$, so $x \in E$.

Theorem 2.7 Let $U \subset \mathbb{R}^m$ be open and path connected, and $f : U \to \mathbb{R}^n$ be differentiable with $Df = 0$ on U. Then f is constant.

Proof

We can Assume (considering the components of f) that $n = 1$. Let $a, b \in U$. By hypothesis, $\exists \gamma : [0,1] \to U$, continuous such that $\gamma(0) = a$, $\gamma(1) = b$. Consider $g = f \circ \gamma : [0,1] \to \mathbb{R}$, which is continuous since f, γ are. Let $s \in [0,1]$, $c = \gamma(s) \in U$. Choose $\varepsilon > 0$ such that $B(c, \varepsilon) \subset U$. Then the corollary to Theorem 2.5 implies: $\forall x \in B(c, \varepsilon)$, $f(x) = f(c)$. γ is continuous, so $\exists \delta > 0$ such that $(s-\delta, s+\delta) \in (0,1)$ and $\forall t \in (s-\delta, s+\delta)$ $\|\gamma(t) - \gamma(s)\| < \varepsilon$, i.e. $\gamma(t) \in B(c, \varepsilon) \Rightarrow f(\gamma(t)) = f(c)$

i.e. g is constant on this interval, so g is differentiable on $(0,1)$ with zero derivative. By the ~~Intermediate~~ Mean Value Theorem, g is constant, i.e. $g(0) = f(a) = g(1) = f(b)$.

So f is constant.

Theorem 28 Let $U \subset \mathbb{R}^m$ be a non-empty open subset. Then U is path connected iff whenever $U_0 \cup U_1 = U$ for open $U_0, U_1 \subset \mathbb{R}^m$ with $U_0 \cap U_1 = \emptyset$, then one of U_0 or U_1 is empty.

Sketch Proof

Suppose $U = U_0 \cup U_1$, with U_0, U_1 disjoint open subsets (non-empty) of $\mathbb{R}^m$

Let $f: U \to \mathbb{R}$ be the function $f(x) = \begin{cases} 0 & x \in U_0 \\ 1 & x \in U_1 \end{cases}$

As U_0, U_1 are open, we can see that f is differentiable with $Df = 0$, contradicting Theorem 27.

Conversely, we observe that the relation on U:

$x \sim y \iff \exists$ continuous $\gamma : [0,1] \to U$ with $\gamma(0) = x$, $\gamma(1) = y$

is an equivalence relation. We can write U as a union of equivalence classes $\{U_i\}$. Each U_i is open, and this is easy to check a U is open. So if there are ~~more~~ at least 2 equivalence classes, then $U = U_0 \cup \bigcup_{i \neq 0} U_i$ is a disjoint union of 2 open subsets.

~~If~~ there is just 1 equivalence class, ~~then~~ iff U is path-connected.

(In general, U_i are called the path-components of U)

Example (Here U is NOT an open subset of $\mathbb{R}^m$)

$U = A \cup B \subset \mathbb{R}^2$, $A = \{(0, y)\}$, $B = \{(x, \sin(\frac{1}{x})) \mid x > 0\}$

This is not path connected, but not the disjoint union of 2 non-empty open sets (A is NOT open in U)

Higher Derivatives

Theorem 29 $U \subset \mathbb{R}^m$ open. $f : U \to \mathbb{R}$ continuously differentiable. i.e. $Df : U \to L(\mathbb{R}^m, \mathbb{R}) \cong \mathbb{R}^m$ is continuous, so $\frac{\partial f}{\partial x_j}$ exist and are continuous on U. Suppose that $\frac{\partial^2 f}{\partial x_i \partial x_j}$ $(1 \leq i, j \leq m)$ exist everywhere in U and are continuous at $x = a \in U$. Then, $\forall i, j$, $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$

11/11/11 Analysis II (15)

Proof Consider one pair (i, j) at a time. We may assume WLOG that $n = 2$. Consider:

$g(h) = \frac{1}{h^2}(f(a_1+h, a_2+h) - f(a_1+h, a_2) - f(a_1, a_2+h) + f(a_1, a_2))$

where $0 < |h| < r$, $B(a, r) \subset U$. $a = (a_1, a_2)$

Applying the Mean Value Theorem to the function

$x \mapsto f(x, a_2+h) - f(x, a_2)$, $\exists b_1$ with $|b_1 - a_1| < |h|$ and

$g(h) = \left[\frac{\partial f}{\partial x_1}(b_1, a_2+h) - \frac{\partial f}{\partial x_1}(b_1, a_2)\right]\frac{1}{h}$

Applying the Mean Value Theorem again, $\exists b_2$ with $|b_2 - a_2| < |h|$ and

$$g(h) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(b_1, b_2) = \frac{\partial}{\partial x_2}\left[\frac{\partial f}{\partial x_1}(b_1, x_2)\right](x_2 = b_2)$$

These 2^{nd} order partial derivatives are continuous at $x = a$,

so $\lim_{h \to 0} g(h) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$

But also $g(h) = \frac{1}{h^2}\left[f(a_1+h, a_2+h) - f(a_1, a_2+h) - f(a_1+h, a_2) + f(a_1, a_2)\right]$

and the same argument shows $\lim_{h \to 0} g(h) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(a)$ □

3

14/11/11

Analysis II (16)

Let $f: U \to \mathbb{R}^n$ ($U \subset \mathbb{R}^m$ open) be differentiable $\quad \cong \mathbb{R}^{mn}$

It's derivative is a function $Df: U \to L(\mathbb{R}^m, \mathbb{R}^n)$ with components $\left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1 \le i \le n \\ 1 \le j \le m}}$. If Df itself is differentiable, its derivative at $a \in U$ will be $D(Df)(a): \mathbb{R}^m \to L(\mathbb{R}^m, \mathbb{R}^n)_{(*)}$, a linear transformation.

From Linear Algebra: Let $\varphi: \mathbb{R}^l \to L(\mathbb{R}^m, \mathbb{R}^l)$ be a linear transformati
(Here $\mathbb{R}^l, \mathbb{R}^m, \mathbb{R}^n$ can be replaced by arbitrary vector spaces.
Define $\Phi: \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}^n$, $\Phi(u, v) = \varphi(u)(v)$ $\quad \varphi(u) \in L(\mathbb{R}^m, \mathbb{R}$
This is not linear but is bilinear.

$\Phi(au + a'u', v) = a\Phi(u, v) + a'\Phi(u', v) \quad \forall u, u' \in \mathbb{R}^l$

$\Phi(u, av + a'v') = a\Phi(u, v) + a'\Phi(u, v') \quad \forall v, v' \in \mathbb{R}^m, \forall a, a' \in \mathbb{R}$

(standard examples) $\mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $(x, y) \mapsto x \cdot y$ scalar product

$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$, $(x, y) \mapsto x \times y$ vector product

$M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to \mathbb{R}$, $(A, B) \mapsto tr(AB)$

$M_n(\mathbb{R}) = \{n \times n \text{ real matrices}\}$

Conversely, if $\Phi: \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}^n$ is bilinear, then it comes from a unique linear map $L(\mathbb{R}^m, \mathbb{R}^n)$, $\varphi(u) = (v \mapsto \Phi(u, v)) \in L(\mathbb{R}^m, \mathbb{R}^n)$
($n = 1$, $\Phi: \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ is said to be a bilinear form)

Definition Suppose Df is differentiable at $a \in U$. Then define

$D^2 f(a): \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$

$D^2 f(a)(u, v) = D(Df)(a)(u)(v)$ $\quad$ (see $(*)$)

From the discussion above, $D^2 f(a)$ is a bilinear map.
(for higher derivatives, it can be more convenient to write $D_a f$ instead of $Df(a)$, i.e. $D_a^2 f(u, v) = D_a(Df)(u)(v)$)

Corollary to Theorem 29: Suppose $U \subset \mathbb{R}^m$ is open, $f: U \to \mathbb{R}^n$ is continuously differentiable, and the 2^{nd} partial derivatives $\frac{\partial^2 f_i}{\partial x_j \partial x_k}$ all exist and are continuous on U. Then $\forall a \in U$, $D_a^2 f$ exists and is a symmetric bilinear form: $D_a^2 f(u, v) = D_a^2 f(v, u)$

Moreover, $D_a^2 f(u,v) = \left(\sum_{j=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 f_i}{\partial x_j \partial x_k}(a)\, u_j v_k \right)_{1\le i\le n} \in \mathbb{R}^n$

[the i^{th} component is $u^T \left(\frac{\partial^2 f_i}{\partial x_j \partial x_k}(a) \right) v$]

Linear Algebra again: Let $B: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ be bilinear and symmetric, $B(u,v) = B(v,u) = \sum_{1\le i,j\le m} b_{ij} u_i v_j$ for a symmetric matrix (b_{ij}). Associated to B is the function:
$Q: \mathbb{R}^m \to \mathbb{R}$, $Q(u) = B(u,u) = \sum_{1\le i,j\le m} b_{ij} u_i u_j$, a <u>quadratic form</u>. This is a homogeneous polynomial of degree 2 in coordinates (u_i). Since B is symmetric, we can recover B from Q by the formula:

$$Q(u+v) = Q(u) + Q(v) + \underbrace{2B(u,v)}_{\leftarrow B(u,v) + B(v,u)}$$

or explicitly by $b_{ij} = \begin{cases} \frac{1}{2} \times \text{coefficient of } u_i v_j & \text{if } i \ne j \\ \text{coefficient of } u_i^2 & \text{if } i = j \end{cases}$

This is a <u>bijection</u> $\left\{ \begin{matrix} \text{symmetric bilinear forms} \\ B: \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \end{matrix} \right\} \xleftrightarrow{\sim} \left\{ \begin{matrix} \text{Quadratic forms} \\ Q: \mathbb{R}^m \to \mathbb{R} \end{matrix} \right\}$

(Here, we can replace $\mathbb{R}^m$ by <u>any</u> vector space, over a field not of characteristic 2).

If B is $D_a^2 f$, we will write $D_a^2 f[u]^2 = D_a^2 f(u,u)$ for the associated quadratic ~~form~~ map.
Now if $D^2 f: U \to \left\{ \begin{matrix} \text{bilinear maps} \\ B: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n \end{matrix} \right\} = Bil(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n) \cong L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n)) \cong \mathbb{R}^{m^2 n}$
and can ask for its derivative.
More generally, assume all partial derivatives of total degree $\le k$ of the components of f exist and are continuous on U. Then there exist higher derivatives $D^r f$ $(1 \le r \le k)$ with

$$D_a^r f: \underbrace{\mathbb{R}^m \times \dots \times \mathbb{R}^m}_{r \text{ times}} \to \mathbb{R}^n$$

$$D_a^r f(u^{(1)}, \dots, u^{(r)}) = D_a(D^{r-1} f)(u^{(1)})(u^{(2)}, \dots, u^{(r)})$$

which is a <u>symmetric</u>, <u>multilinear</u> mapping (linear in each $u^{(i)}$ and doesn't depend on their order).

14/11/11

Analysis II (16)

We have:

$$D_a^r f(u^{(1)}, \ldots, u^{(r)})_i = \sum_{j_1, j_2, \ldots, j_r = 1}^{m} A_i^{j_1, \ldots, j_r} u^{(1)}_{j_1} \ldots u^{(r)}_{j_r} \quad (1 \le i \le n)$$

i^{th} component

coordinates of $u^{(1)}, \ldots, u^{(r)}$

where $$A_i^{j_1, \ldots, j_r} = \frac{\partial^r f_i}{\partial x_{j_1} \ldots \partial x_{j_r}}(a)$$

Write $D_a^r f[u]^r = D_a^r f(u, \ldots, u)$
which is a vector of homogeneous polynomials of degree r in the coordinates $\{u_j\}$.

Just as for bilinear maps, from the map $u \mapsto D_a^r f[u]^r$ ($\mathbb{R}^m \to \mathbb{R}^n$), you can recover the full multilinear map $D_a^r f : \underbrace{\mathbb{R}^m \times \mathbb{R}^m \times \ldots \times \mathbb{R}^m}_{r \text{ copies}} \to \mathbb{R}^n$ by an explicit formula that we don't need explicitly (polarisation).

18/11/11 Analysis II (17)

Remarks

Let W be a finite dimensional real vector space, $\dim W = n$

$W \cong \mathbb{R}^n$ as a vector space. Any two norms on W are equivalent.

$U \underset{\text{open}}{\subseteq} \mathbb{R}^m$, $f : U \to W$ any function.

$a \in$ The notion of derivative of f at a makes sense that it will be a linear map $A = D_a f : \mathbb{R}^m \to W$ such that

$\frac{1}{\|h\|}\left(f(a+h) - f(a) - A(h)\right) \to 0$ as $h \to 0$

where "$\to 0$" means tends to limit 0 with respect to some (or any) norm on W (all norms being equivalent means that convergence for 1 norm $\Rightarrow$ convergence for any norm)

[We could also consider $U \underset{\text{open}}{\subseteq} V$, V some finite dimensional vector space, if desired, then $D_a f$ would be an element of $L(V, W)$.

So if $\mathbb{R}^m \supset U \xrightarrow{f} \mathbb{R}^n$ is differentiable,

Df is a function from U to $L(\mathbb{R}^m, \mathbb{R}^n) = V$, say

$\therefore D_a(Df)$ is a linear map $\mathbb{R}^m \to V$

Last time we defined higher derivatives of $f : U \to \mathbb{R}^n$ ($U \subset \mathbb{R}^m$)

$D_a^k f : \underbrace{\mathbb{R}^m \times \dots \times \mathbb{R}^m}_{k \text{ times}} \to \mathbb{R}^n$, given by the order k partial derivatives of components of f,

and $D_a^k f[u]^k = D_a^k f(u, \dots, u) \in \mathbb{R}^n$, $u \in \mathbb{R}^m$,

which is a polynomial function from $\mathbb{R}^m$ to $\mathbb{R}^n$, homogeneous of degree k.

Another interpretation (or if you like, definition) of $D_a^k f[u]^k$ is in terms of higher order directional derivatives.

Suppose $n = 1$, $a \in U$, $h \in \mathbb{R}^m$ such that $[a, a+h] = \{a + th \mid t \in [0,1]\} \subset U$

Let $\gamma(t) = a + th$ $(t \in [0,1])$ and $g = f \circ \gamma$, $g(t) = f(a+th)$

Assuming that the derivatives exist:

$g'(t) = D_{\gamma(t)} f \circ D_t \gamma = D_{a+th} f(h)$ $(D_{a+th} f : \mathbb{R}^m \to \mathbb{R})$

Then $g''(t) = D_{a+th}(Df(h)) \circ \overset{=h}{\gamma'(t)} = D^2_{a+th} f(h,h) = D^2_{a+th} f[h]^2$

(Chain rule again)

and likewise $g^{(k)}(t) = D^k_{a+th} f[h]^k$
In particular : $D_a{}^k f[h]^k = g^{(k)}(0) \leftarrow$ (*)

Theorem 30 (Taylor's Theorem) $\mathbb{R}^m \underset{open}{\supset} U$, $f: U \to \mathbb{R}$, k times differentiable. Then $\forall a \in U$, $h \in \mathbb{R}^m$ such that $[a, a+h] \subset U$,

$$f(a+h) = f(a) + D_a f[h] + \frac{1}{2!} D_a^2 f[h]^2 + \frac{1}{3!} D_a^3 f[h]^3 + \dots$$
$$\dots + \frac{1}{(k-1)!} D_a^{k-1} f[h]^{k-1} + \frac{1}{k!} D_b^k f[h]^k$$

for some $b \in [a, a+h]$

Proof

Let $g(t) = f(a+th)$, $t \in (-\varepsilon, 1+\varepsilon)$ for $\varepsilon > 0$ chosen so that $\{a+th \mid t \in (-\varepsilon, 1+\varepsilon)\} \subset U$. Then by Taylor's Theorem in 1 variable:

$$g(t) = \sum_{p=0}^{k-1} \frac{1}{p!} g^{(p)}(0) t^p + \frac{1}{k!} g^{(k)}(s) t^k, \quad s \in [0, t]$$

and then apply (*)

$$f(a+ht) = \sum_{p=0}^{k-1} \frac{1}{p!} D_a^p f[h]^p t^p + \frac{1}{k!} D^k_{a+sh} f[h]^k t^k$$

Set $t = 1$ and $b = a + sh$ to get the desired result.

Note that this doesn't apply directly to $f: U \to \mathbb{R}^n$, $n > 1$ as we will have different b's for each component of f. However, as with the mean-value inequality, we can prove :

Corollary Suppose $f: U \to \mathbb{R}^n$ (U open $\subseteq \mathbb{R}^m$) has derivatives of all orders ($\Leftrightarrow$ all partial derivatives of all orders exist and are continuous on U). We say that f is infinitely differentiable on C^∞ (or smooth)
If ~~when~~ $a \in B(a, r) \subset U$ and $k \geq 0$, then $\forall h$ with $\|h\| < r$

$$f(a+h) = \sum_{p=0}^{k} \frac{1}{p!} D_a^p f[h]^p + R_k(h) \quad \text{where } \frac{R(h)}{\|h\|^k} \to 0 \text{ as } h \to 0$$

(actually $\|R_k(h)\| < c\|h\|^{k+1}$ as $h \to 0$)

18/11/11

Analysis II (18)

5 Metric Spaces

Definition A *metric* on a set X is a map $d : X \times X \to \mathbb{R}$ (the *metric* or *distance* function) such that

i) $d(x, y) = d(y, x)$

ii) $d(x, y) = 0 \iff x = y$

iii) ("triangle inequality") $d(x, z) \leq d(x, y) + d(y, z)$

Examples
- $\mathbb{R}^n$, Euclidean distance $d(x, y) = \left(\sum (x_i - y_i)^2\right)^{\frac{1}{2}}$
- $(V, \|\cdot\|)$ any normed vector space, $d(x, y) = \|x - y\|$ (our axioms are then equivalent to the axioms for norms)
- Any subset of a normed space with some metric

Extreme / trivial / artificial metrics
- X any set : $d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & x \neq y \end{cases}$ "discrete metric"
- "tape measure metric" (X, d) any metric space
  (X, d') $d'(x, y) = \min(d(x, y), 1000)$
- $X = \{\text{finite subsets of } \mathbb{N}\} \ni S, T$
  Define the symmetric ~~distance~~ difference of S, T to be
  $S \triangle T = S \cup T \setminus (S \cap T)$
  $d(S, T) = |S \triangle T|$
- BR metric on $\mathbb{R}^2$: $d(x, y) = \begin{cases} 0 & \text{if } x = y \\ \|x\| + \|y\| & \text{if } x \neq y \end{cases}$

Two metrics d, d' on X are said to be *Lipschitz equivalent* if $\exists A, B > 0$ such that $\forall x, y \in X$

$$A\, d(x, y) \leq d'(x, y) \leq B\, d(x, y)$$

(an equivalence relation on metrics)

Two norms $\|\cdot\|, \|\cdot\|'$ on a vector space V are equivalent iff the corresponding metrics are Lipschitz equivalent (by definition).

So any metric on $\mathbb{R}^n$ coming from a norm is Lipschitz equivalent to Euclidean metric (as any 2 norms on $\mathbb{R}^n$ are equivalent). But there are many other metrics on $\mathbb{R}^n$.

(For example, the usual metric and the tapemeasure metric on $\mathbb{R}$ are not equivalent).
For all practical purposes, Lipschitz equivalent metrics are indistinguishable

(N.B a metric space is a pair (X, d) with X a set, d a metric on X)

Metric subspaces
(X, d) a metric space, $Y \subset X$ any subset. Then $d|_{Y \times Y} : Y \times Y \to \mathbb{R}$ (the restriction of d to $Y \times Y$) is a metric on Y, called the induced or subspace metric on Y, and we say Y (with this metric) is a metric subspace of (X, d)

Products $(X, d), (X', d')$ metric spaces
Consider the product $X \times X' = \{(x, x') \mid x \in X, x' \in X'\}$
There are at least 3 ways to put a metric on $X \times X'$.

$$\begin{cases} d_\infty((x,x'),(y,y')) \mapsto \max\{d(x,y), d'(x',y')\} \\ d_1 \quad d(x,y) + d'(x',y') \\ d_2 \quad \sqrt{d(x,y)^2 + d'(x',y')^2} \end{cases}$$

$d_\infty \leq d_2 < d_1 \leq 2 d_\infty$ so these are all equivalent metrics on $X \times X'$

"Metrics in Geometry"
Topology - About Shape Geometry - About (metric) distance
Spherical Geometry $X = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid \|x\| = 1\}$ (written S^2)
($\|x\| < 1$ ball)

A great circle on X is any subset of the form $C = X \cap \Pi$, where Π is a plane (linear subspace of dimension 2) through O.
If $x, y \in X$, then $\exists$ a great circle containing them (only one unless $y \in \{x, -x\}$)

Spherical distance $d(x, y)$ = arc length from x to y along the great circle

This is a metric on X.

18/11/11

Analysis II (18)

Another description: $\gamma : [a,b] \to X$ continuously differentiable.
$L(\gamma) = \int_a^b \|\gamma'(t)\| \, dt$ (arc length of γ)

Then $d(x,y) = \inf \{ L(\gamma) \mid \gamma : [a,b] \to X$ continuously differentiable, $\gamma(a) = x, \gamma(b) = y \}$

It is easy to verify that spherical distance is Lipschitz equivalent to Euclidean distance on X.

Another example is hyperbolic geometry in the "hyperbolic plane"
D = open unit disc (a ball) in $\mathbb{R}^2 \cong \mathbb{C}$
$\gamma : [a,b] \to D$ continuously differentiable
We define the hyperbolic length of γ to be:
~~$L_{hyp}(\gamma) = \int_a^b$~~

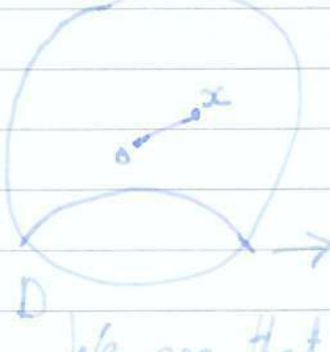
$$L_{hyp}(\gamma) = \int_a^b \frac{\|\gamma'(t)\|}{1 - \|\gamma(t)\|^2} \, dt$$

$d_{hyp}(x,y) = \inf \{ L_{hyp}(\gamma) \mid \gamma : [a,b] \to D, \gamma(a) = x, \gamma(b) = y \}$ for $x, y \in D$.

(As points get closer to the boundary circle, they get much further apart).

$\|x\|_{Eucl} = r$

$$d_{hyp}(0,x) = \int_0^r 2 \frac{dt}{1-t^2} = 2\tanh^{-1} r \to \infty \text{ as } r \to 1$$



→ Paths of shortest hyperbolic distance are ~~arcs~~ circular paths meeting the boundary orthogonally.

We see that d_{hyp} is very far from being Lipschitz equivalent to the Euclidean norm.

21/11/11

Analysis II (19)

Definition (X,d), (X',d') metric spaces
A mapping $f: X \to X'$ is <u>continuous</u> if $\forall x \in X, \forall \varepsilon > 0$ $\exists \delta > 0$ such that $\forall y \in X$, $d(x,y) < \delta \Rightarrow d'(f(x), f(y)) < \varepsilon$
and <u>uniformly continuous</u> if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\forall x, y \in X$, $x,y < \delta \Rightarrow d'(f(x), f(y)) < \varepsilon$.
A map is Lipschitz if $\exists k \geq 0$ such that $\forall x, y \in X$,
$d'(f(x), f(y)) \leq k\, d(x,y)$ (with k called the <u>Lipschitz const</u> of f)

f Lipschitz $\underset{(\delta < \frac{\varepsilon}{k})}{\Rightarrow}$ f uniformly continuous $\Rightarrow$ f continuous

These Notions are useful even for $X = X'$, f the identity map.

<u>Example</u>
$X = X' = \mathbb{R}$, d' = Euclidean metric $|x-y|$, $d(x,y) = \min\{1, |x-y\}$
f = identity mapping
Then f <u>is</u> uniformly continuous, but <u>not</u> Lipschitz.

Topology of Metric Spaces

Just as for normed spaces, the metric determines open sets.
<u>Definition</u> (X,d) a metric space, $a \in X$, $r > 0$
$B(a,r) = \{x \in X, d(x,a) < r\}$ open ball of radius r
$U \subset X$ is <u>open</u> if $\forall a \in U$, $\exists r > 0$ such that $B(a,r) \subset U$
$E \subset X$ is <u>closed</u> if $X \setminus E$ is an <u>open</u> subset of X.
<u>Remark</u>
Both of these notions are <u>relative</u> to the ambient (containing) space X.

<u>Example</u> $X = \mathbb{R}$, $d(x,y) = |x-y|$, $Y = [0,1)$ with induced met
Y is both open and closed as a subset of Y
e.g. in Y $B(0,r) = [0,r)$ $(r<1)$
but neither open nor closed as a subset of X.

<u>Remark</u> if $a \in X$, $r > 0$ then the ball $B(a,r)$ is an open subset of X,

[Proof: let $x \in B(a,r)$; let $r' = d(a,x) < r$.
Put $\varepsilon = r - r'$. Then $y \in B(x,\varepsilon) \subset X$, $d(y,a) \leq d(x,a) + d(x,$
$d(y,a) < \varepsilon + r' \leq r$ i.e. $y \in B(a,r)$
So $B(x,\varepsilon) \subset B(a,r)$

<u>Remark</u> Open balls can be closed
$X = \mathbb{Z}$, $d(x,y) = |x-y|$
$B(0,r) = \{n \in \mathbb{Z} \mid |n| < r\}$
Every subset of X is open because $B(a,\frac{1}{2}) = \{a\}$ if $a \in \mathbb{Z}$.
(so every subset is closed). $B(0,4) = B(0,\pi)$

<u>Properties of open and closed sets</u>
If $\{U_i\}$ is a collection (finite or infinite) of open subsets of X, then $\bigcup U_i$ is also an open <u>subset</u>. If $\{U_i\}$ is finite, then the intersection of the U_i is also open. This does not hold for infinite intersections.

<u>Example</u> $X = \mathbb{R}$, $U_i = B(0,\frac{1}{i}) = (-\frac{1}{i},\frac{1}{i})$ is open in X $\forall i \geq$
$\bigcap U_i = \{0\}$, <u>not</u> open in $\mathbb{R}$ (with Euclidean metric)
So an infinite intersection of open subsets is not necessarily open.

For closed subsets $\{E_i\}$ any intersection is closed and finite unions are closed.
For any x, the <u>singleton</u> subset $\{x\}$ is closed (since if $y \neq x$, $B(y, d(x,y)) \subset X \setminus \{x\}$ so the complement is open.
Any <u>finite</u> subset of X is closed.

The set of open subsets of a metric space X is called the <u>topology</u> of X. A property of X which depends only its open sets (i.e. on its topology) is said to be <u>topology</u> (examples: continuity; limits)

$E \subset X$ any subset. $x \in X$ is a <u>limit point</u> of E if $\forall \varepsilon > 0$, $\exists y \in E$ with $0 < d(x,y) < \varepsilon$

21/11/11 Analysis II (19)

Proposition 31

$E \subset X$ is closed $\iff$ E contains all of its limit points.

[Proof: Same as for normed spaces, Theorem 13 i)]

So we get the same notion of closed sets as for normed spaces.

Useful Terminology X a metric space, $x \in X$

An open neighbourhood of x in X is any open subset of X containing x.

A neighbourhood of x in X is any subset $E \subset X$ such that $\exists r > 0$ with $B(x, r) \subset E$. (In particular, an open neighbourhood of X is also a neighbourhood of x)

(some authors insist that neighbourhoods are open)

Limits (X, d) a metric space. $x, (x_n)_{n \in \mathbb{N}} \in X$.

We say that $(x_n) \to x$ or $x = \lim\limits_{n \to \infty} x_n$ if $\forall \varepsilon > 0$, $\exists N = N_\varepsilon$ such that $n \geq N \Rightarrow d(x, x_n) < \varepsilon$

Equivalently $(x_n) \to x \iff (d(x, x_n))_n \to 0$

Proposition 32 $(x_n) \to x$

$\iff$ For any neighbourhood V of x in X, then $x_n \in V$ for all but finitely many n.

Proof

If $(x_n) \to x$ and V is a neighbourhood of x, then $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset V$. Also $\exists N = N_\varepsilon$ such that $\forall n \geq N$, $d(x_n, x) < \varepsilon$, so that $\forall n \geq N$, $x_n \in V$.

The converse is clear because any $B(x, \varepsilon)$ is a neighbourhood of x

Consequently, convergence depends only on the topology of X.

23/11/11 Analysis II (20)

The notion of convergence is "topological".

Theorem 33 (X,d), (X',d') metric spaces $f: X \to X'$

The following are equivalent:

i) f is continuous ii) $\forall x_n$ $(n \in \mathbb{N})$ with $(x_n) \to x$, we have $f(x_n) \to f($

iii) $\forall V$ open in X', the inverse image $f^{-1}(V) = \{x \in X \mid f(x) \in V\}$ is open in X

(iii') $\forall$ closed $E \subset X'$, $f^{-1}(E)$ is closed in X.

Proof (i) $\Leftrightarrow$ (ii) the same as the proof of Theorem 15

Since $f^{-1}(X' \setminus E) = X \setminus f^{-1}(E)$, (iii) $\Leftrightarrow$ (iii')

(i) $\Rightarrow$ (iii) Let $V \subset X'$ and let $U = f^{-1}(V) \subset X$, $x \in U$. We wish to show that U is open. As V is open, $\exists \varepsilon > 0$ such that $B(f(x), \varepsilon) \subset V$.

As f is continuous, $\exists \delta > 0$ such that $\forall y \in B(x, \delta)$, $f(y) \in B(f(x), \varepsilon$

So $B(x,\delta) \subset U$. $\therefore$ U is open.

(iii) $\Rightarrow$ (i) Let $x \in X$, $\varepsilon > 0$. Consider $V = B(f(x), \varepsilon) \subset X'$. $f^{-1}(V)$ is open by hypothesis, so $\exists \delta > 0$ such that $B(x, \delta) \subset f^{-1}(V)$

This means that $d(y, x) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$

i.e. f is continuous at x.

[This shows that continuity is a topological property. By (iii) it depends only on the open sets of X, X']

Tautological Example (X, d), $d: X \times X \to \mathbb{R}$ (with Euclidean metric on $\mathbb{R}$) is continuous (follows easily from the definition)

Other easy general examples:

① $f: X \times Y \to Z$ continuous. Then $\forall y \in Y$, the map
$f(-, y): X \to Z$
$x \mapsto f(x,y)$ is also continuous

$$x \longmapsto (x,y)$$
$$X \to X \times Y \xrightarrow{f} Z$$

② Composites of continuous maps are continuous (① is a special case)

Analysis on metric spaces

Definition: (X,d) is a metric space, (x_n) a sequence in X. We say that (x_n) is a Cauchy Sequence if $\forall \varepsilon > 0$, $\exists N = N_\varepsilon$ such that $\forall m,n \geq N$, $d(x_m, x_n) < \varepsilon$

Theorem 34 $f: X \to X'$ uniformly continuous, (x_n) a Cauchy Sequence in X. Then $(f(x_n))$ is a Cauchy sequence in X'.

Example Let $X = [0,1)$, $X' = \mathbb{R}^+$ $\to \mathbb{R}_{>0}$, both with Euclidean metric

$f: X \to X'$, $f(x) = \frac{x}{1-x}$ is continuous, bijective and has a continuous inverse. f is not uniformly continuous (easy to show).

is a homeomorphism $\to$ $\forall U \subset X$, U is open iff $f(U)$ is open.

The sequence $x_n = 1 - \frac{1}{n}$ $(n \geq 2)$ is a Cauchy sequence in X. But $f(x_n) = n-1$ is NOT a Cauchy sequence in X'.

Proof

f is uniformly continuous, so if $\varepsilon > 0$, $\exists \delta > 0$ such that $\forall x,y \in X$, $d(x,y) < \delta \Rightarrow d'(f(x), f(y)) < \varepsilon$

(x_n) a Cauchy Sequence $\Rightarrow \exists N = N_\delta$ such that $\forall m,n \geq N$, $d'(x_m, x_n) <$

$\Rightarrow d'(f(x_m), f(x_n)) < \varepsilon$, i.e. $(f(x_n))$ is Cauchy.

Proposition 35

i) If $(x_n) \to x$ in X, then (x_n) is Cauchy.

ii) If (x_n) is a Cauchy Sequence in X and has a convergent subsequence, then (x_n) converges in X.

23/11/11

Analysis II (20)

<u>Proof</u>

i) Same as Proposition 9

ii) Let $(x_{\sigma(n)})$ be a convergent subsequence, $(x_{\sigma(n)}) \to x \in X$

($\sigma : \mathbb{N} \to \mathbb{N}$ a strictly increasing function)

$\sigma(n) \geq n$, so $d(x_n, x_{\sigma(n)}) \to 0$ as (x_n) is Cauchy.

Also, $d(x_{\sigma(n)}, x) \to 0$ by convergence. So by the Triangle Inequality $d(x_n, x) \to 0$, i.e. $(x_n) \to x$ $\Box$

Just as for normed spaces:

<u>Definition</u> (X, d) is complete if every Cauchy Sequence in X is convergent (in X, of course).

<u>Example</u> $X = \mathbb{R}^n$ is complete for the Euclidean metric.

$X = (0,1)$ with the Euclidean metric is NOT complete.

$x_n = 1 - \frac{1}{n}$ is a Cauchy sequence which doesn't converge in X.

When is a metric space complete?

Some useful facts:

<u>Proposition 36</u> X a complete metric space. Let $Y \subset X$. When is Y complete

Y (with the induced metric) is complete $\Leftrightarrow$ Y is closed.

<u>Proof</u> Let (y_n) be any sequence in Y. Then (y_n) is a Cauchy sequence in Y (for the induced metric) iff it is a Cauchy sequence in X, iff (y_n) is convergent in X.

If Y is closed, then it contains all of its limit points, so this holds iff (y_n) converges in Y i.e. Y is complete.

If Y is <u>not</u> closed, $\exists x \in X \setminus Y$, x a limit point of Y. Let (y_n) be a sequence in Y with $(y_n) \to x$. ~~So (y_n) converges~~

So (y_n) is a Cauchy sequence in Y which doesn't converge, so Y is not complete.

24/11/11 Analysis II (21)

Example

$C[a,b] = \{\text{continuous functions } [a,b] \to \mathbb{R}\}$, $d(f,g) = \sup_{[a,b]} |f-g|$ is complete (by the Cauchy criterion for uniform convergence)

$X = \{f \in C[a,b] \mid \forall x \in [a,b], |f(x)| \leq R\} = \{f \mid \|f\|_\infty \leq R\}$ is closed in $C[a,b]$ (check). So X is complete.

What about Bolzano Weierstrass in the context of metric spaces? (See Met + Top.)

Definition (X,d) is sequentially compact if every sequence in X has a convergent subsequence. [This depends only on convergence so is a topological property. For topological spaces there is another definition of compactness using open covers; these are equivalent for metric spaces]

Example

$X =$ A closed bounded subset of $\mathbb{R}^n$. X is compact by Bolzano-Weierstrass

If X (X,d) is compact then:

- X is complete (every Cauchy sequence has a convergent subsequence, hence by Proposition 35(ii) converges)
- X is bounded $(\exists x_0 \in X, R > 0$ such that $B(x_0, R) = X)$ (if not, then take any x_0, then $\exists x_n \in X$ with $d(x_n, x_0) > n$ and x_n doesn't converge)

Those two conditions are not sufficient.

Example 1 Closed unit ball in ℓ_∞ (see end of lecture 7, remark after Theorem 1(

Example 2 $X = \mathbb{R}$, $d(x,y) = \min(1, |x-y|)$, $X = B(0,R)$ for any $R > 1$. X is complete (easy to see). But $(x_n) = (n)$ has no convergent subsequence.

<u>Example 3</u> $X = \{ f \in C[0,1] \mid \|f\|_{sup} \leq 1 \}$ is bounded and complete

f_n 

$\forall n \neq m$, $\|f_n - f_m\| = 1$. So (f_n) doesn't contain a convergent subsequence.

<u>Definition</u> We say that (X,d) is <u>totally bounded</u> if $\forall \varepsilon > 0$, $\exists$ a finite subset $\{x_1, \dots, x_N\} \subset X$, $(N = N_\varepsilon)$ such that $X = \bigcup_{i=1}^{N} B(x_i, \varepsilon)$

<u>Theorem</u> (X,d) a metric space. → depend on the metric

Topological Property → X is compact $\iff$ X is complete and totally bounded

(Exercise: Check that examples 1 to 3 are <u>not</u> totally bounded)

<u>5 Contraction Mapping Theorem</u>

Example of a "fixed point theorem": Let $f: X \to X$. When does there exist $x \in X$ with $x = f(x)$? (i.e. a fixed point of f)

<u>Definition</u> (X,d) a metric space

A <u>contraction mapping</u> on (X,d) is a mapping $f: X \to X$ such that $\exists \lambda \in [0,1)$ such that $\forall x, y \in X$, $d(f(x), f(y)) \leq \lambda d(x,y)$

[i.e. f is Lipschitz with constant $\lambda < 1$. So in particular, f is continuous]

<u>Remark</u> The "more natural" condition would appear to be $\forall x, y \in X$, $d(f(x), f(y)) < d(x,y)$ but <u>uniformity</u> (having a fixed $\lambda < 1$ is crucial).

<u>Examples</u> $X = \mathbb{R}$, $f(x) = ax + b$, f is a contraction $\iff |a| = \lambda < 1$

$X = \mathbb{R}^n$, $f(x) = Ax + b$, $b \in \mathbb{R}^n$, $A \in L(\mathbb{R}^n, \mathbb{R}^n)$

f is a contraction $\iff \|A\| < 1$ (operator norm)

25/11/11 Analysis II 21

$X = (0, \frac{\pi}{2})$, $f(x) = \sin(x)$

Mean Value Theorem $\Rightarrow |f(y) - f(x)| = |y-x|\cos t$, $x < t < y$

$|f(x) - f(y)| < |y - x|$. But this is not a contraction as

$\sup_t \{\cos t\} = 1$

Theorem 37 (Contraction Mapping Theorem) X a non-empty metric space $f : X \to X$ a contraction. If X is complete, then f has a fixed point, which is unique.

Proof Suppose $\forall x, y \in X$, $d(f(x), f(y)) \leq \lambda d(x, y)$, $0 \leq \lambda < 1$

Uniqueness: If $f(x) = x$, $f(y) = y$, then $d(x,y) \leq \lambda d(x,y)$, but $\lambda < 1$

$\Rightarrow d(x,y) = 0$, $x = y$

Existence Pick $x_0 \in X$, and let $x_{n+1} = f(x_n)$ $\forall n \geq 0$. We claim that (x_n) is a Cauchy Sequence. $d(x_{n+1}, x_n) \leq \lambda d(x_n, x_{n-1})$

$\Rightarrow d(x_{n+1}, x_n) \leq \lambda^n d(x_1, x_0)$

So if $m \geq n$, $d(x_m, x_n) \leq d(x_m, x_{m-1}) + \dots + d(x_{n+1}, x_n)$

$d(x_m, x_n) \leq (\lambda^{m-1} + \dots + \lambda^n) d(x_1, x_0) \leq \frac{\lambda^n}{1-\lambda} d(x_1, x_0)$

So (x_n) is Cauchy. As X is complete, $(x_n) \to x \in X$.

$$d(f(x), x) \leq \underbrace{d(f(x), f(x_n))}_{\to 0,\ f \text{ continuous},\ x_n \to x} + d(\underbrace{f(x_n)}_{= x_{n+1}}, x_n) + \underbrace{d(x_n, x)}_{x_n \to x}$$

$\Rightarrow d(f(x), x) = 0$, $f(x) = x$

Remark Completeness is essential. $f : (0,1) \to (0,1)$, $f(x) = \frac{x}{2}$

Example $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $b \in \mathbb{R}^n$, $\|A\| = \lambda < 1$

Apply our theorem to $f(x) = Ax + b$ $\Rightarrow \exists!\, x \in \mathbb{R}^n$ (unique) with $f(x) = x$

i.e. $(I - A)x = b$ (can check directly that $\forall x$, $\|Ax\| \leq \lambda \|x\|$ means that 1 is not an eigenvalue of A)

28/11/11

Analysis II (28)

Contraction Mapping Theorem

Corollary (X, d) complete, $f: X \to X$

Assume that for some $m \geq 1$, $f^m = \underbrace{f \circ \ldots \circ f}^{m \text{ times}}$ is a contraction.

Then f has a unique fixed point.

N.B f needn't be a contraction! (Exercise: find an example)

Proof Any fixed point of f is a fixed point of f^m

Contraction Mapping Theorem $\Rightarrow$ f^m has a unique fixed point $f^m(x) = x \in X$

But then $f^m(f(x)) = f(f^m(x)) = f(x)$ so $f(x)$ is a fixed point of f^m. So $f(x) = x$

Two Applications of the Contraction Mapping Theorem — In Analysis

(A) Inverse function Theorem

Let $U \underset{\text{open}}{\subset} \mathbb{R}^n$, $f: U \to \mathbb{R}^n$, differentiable with Df continuous ($f \in C^1$)

Let $a \in U$ with $Df(a) \in L(\mathbb{R}^n, \mathbb{R}^n)$ invertible. Then $\exists V$, an open neighbourhood of a in U such that

i) $f(V) \subset \mathbb{R}^n$ is open and $f|_V : V \to f(V)$ is a bijection

ii) The inverse (which exists by (i)) $f|_V^{-1} : f(V) \to V \subset \mathbb{R}^n$ is C^1 and $D(f|_V^{-1})(f(a)) = (Df)(a)^{-1}$

Remarks The last part of (ii) follows from the rest and the chain rule.

Assuming $g = f|_V^{-1} : f(V) \to V$ exists,

$g \circ f = id_V : x \mapsto x$, so $Dg(f(x)) \cdot Df(x) = I$ ← derivative of id_V

When $n=1$ $f: U \to \mathbb{R}$, $a \in U$ such that $f'(a) \neq 0$ (>0 WLOG)

f' continuous $\Rightarrow \exists \varepsilon > 0$ such that $f' > 0$ on $(a-\varepsilon, a+\varepsilon)$

Hence (Mean Value Theorem) f is monotone, strictly increasing on $(a-\varepsilon, a+\varepsilon)$

So $f:(a-\varepsilon, a+\varepsilon) \to f\big((a-\varepsilon, a+\varepsilon)\big)$ is bijective. The rest is easy.
(This is easy because $\mathbb{R}$ is <u>ordered</u>)

<u>Proof</u> Let $A = Df(a)$. Replacing f by $\hat{f}(x) = f(A^{-1}(x)+a) - f(a)$ *

We may assume WLOG that $a = 0 \in \mathcal{U}$, $Df(0) = I$

As Df is continuous on $\mathcal{U}$, $\exists\, r > 0$ such that

$\bar{B}(0,2r) = \{x \in \mathbb{R}^n \mid \|x\| \leq 2r\} \subset \mathcal{U}$ closed ball

and $\forall x \in \bar{B}(0,2r)$, $\|Df(x) - I\| \leq \frac{1}{2}$ (1) operator norm

(which implies that Df is invertible on $\bar{B}(0,2r)$ by example after Contraction Mapping Theorem)

We will find $V \subset \mathcal{U}$ with $0 \in V$, $f: V \to B(0,r)$ bijective.

For $b \in B(0,r)$, set $T_b: \mathcal{U} \to \mathbb{R}^n$, $T_b(x) = \underbrace{(I-f)(x) + b}_{= x - f(x) + b}$

Note: $T_b(x) = x \iff f(x) = b$ and $DT_b(x) = I - Df(x)$

Apply the Mean Value Inequality (Theorem 25) using (1) to get:

$\forall x, y \in \bar{B}(0,2r)$, $\|T_b(x) - T_b(y)\| \leq \frac{1}{2}\|x-y\|$ (2)

Also $\|T_b(x)\| \leq \|T_b(x) - T_b(0)\| + \|b\|$ $(b = T_b(0))$

$\leq \frac{1}{2}\|x\| + \|b\|$

If $\|x\| \leq 2r$, $\|T_b(x)\| < 2r$ (by (2))

i.e. $T_b: \bar{B}(0,2r) \to B(0,2r) \subset \bar{B}(0,2r)$ and is a contraction on $\bar{B}(0,2r) = X$ by (2). As X is complete, we therefore deduce that T_b has a unique fixed point c, necessarily in $B(0,2r)$ so $f(c) = b$.

So if $W = B(0,r)$, $V = B(0,2r) \cap f^{-1}(W)$, then

$\forall b \in W$, $\exists!\, c \in V$ such that $f(c) = b$.

So $f|_V: V \to W$ is <u>bijective</u>. Let g be its inverse.

28/11/11 Analysis II (22)

If $b_1, b_2 \in W$, $a_i = g(b_i) \in V$, then

$$\|a_1 - a_2\| = \|T_{b_1}(a_1) - T_{b_2}(a_2)\|$$
$$\leq \|(I-f)(a_1) - (I-f)(a_2)\| + \|b_1 - b_2\|$$
$$= \|T_{b_1}(a_1) - T_{b_1}(a_2)\| + \|b_1 - b_2\| \leq \tfrac{1}{2}\|a_1 - a_2\| + \|b_1 - b_2\| \quad \text{by } (2)$$

$$\|a_1 - a_2\| \leq 2\|b_1 - b_2\| \qquad (3)$$

So g is Lipschitz and hence continuous.

Finally, we show that g is differentiable at $b = f(a) \in W$.

Changing variables again, we reduce to the case $a = 0$, $Df(0) = I$

(Since by (1) we know that $Df(a)$ is invertible)

Let $k \in B(0, \varepsilon) \subset W$. Then $\underset{f(h) = k}{h = g(k)} \in V$, $k = f(h)$.

$$\text{So, } \frac{\|g(k) - k\|}{\|k\|} = \frac{\|h - f(h)\|}{\|k\|} = \frac{o(h)}{\|k\|} \quad \text{as } Df(0) = I$$
$$= \frac{o(k)}{k} \text{ by } (3)$$
$$\to 0 \text{ as } k \to 0$$

So g is differentiable, and as we saw, its derivative is $(Df)^{-1}$, hence is continuous.

28/11/11

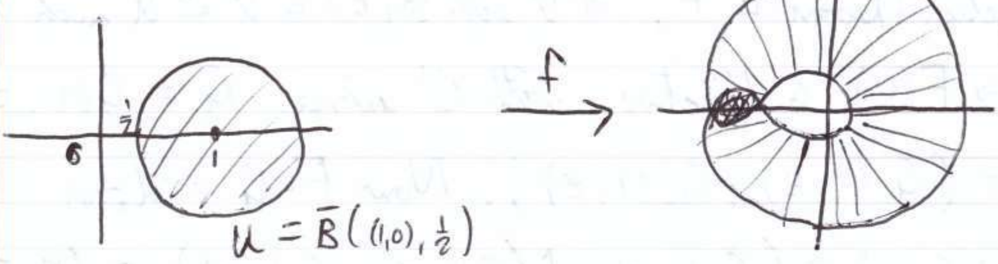
Analysis II (23)

<u>Remark</u> 1-variable case, $f : \mathbb{R} \to \mathbb{R}$ differentiable, $f' \neq 0$ on (a,b) ($f' > 0$ say) then $f : (a,b) \to (f(a), f(b))$ is bijective

In $\mathbb{R}^n$, $n > 1$, it is not necessarily the case that $\mathbb{R}^n \supset U \xrightarrow{f} \mathbb{R}^n$, Df invertible on U implies that f is injective (even if U is <u>connected</u>, to avoid a silly counterexample).

<u>Example</u> $\mathbb{R}^2 \cong \mathbb{C} \xrightarrow{f} \mathbb{C} \cong \mathbb{R}^2$, $f(z) = z^n = u_n(x,y) + i v_n(x,y)$

$(x,y) \mapsto x + iy = z$. It is easy to show that f is differentiable and $Df \neq 0$ except at 0.



$U = \bar{B}((1,0), \tfrac{1}{2})$

Df is invertible on U but $U \to f(U)$ is <u>not</u> Bijective

$y = (y_1, \dots, y_m) = f(x_1, \dots, x_n)$. Df invertible.

$\Rightarrow$ Locally we can write $x = g(y)$.

<u>Implicit Function Theorem</u> $f(x,y) = 0$, when can we "solve" $y = g(x)$.

e.g. $f = x^2 + y^2 - 1$ $\{f = 0\}$

Except at $(\pm 1, 0)$ we can locally solve by $y = \sqrt{1 - x^2}$, a C^1 function of x. At $(1,0)$, any neighbourhood must contain points with y both > 0 and < 0, so no local inverse exists.

<u>Implicit Function Theorem:</u> $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n \underset{\text{open}}{\supset} U \xrightarrow{f} \mathbb{R}^n$

Suppose f is C^1, $(a,b) \in U$, $Df = (D_1 f \mid D_2 f)$, $D_1 f : U \to L(\mathbb{R}^m, \mathbb{R}^n)$, $D_2 f : U \to L(\mathbb{R}^n, \mathbb{R}^n)$

with $f(a,b) = 0$ and $D_2 f$ invertible at (a,b)

~~∃ ∃ V, open~~

Then $\exists$ open neighbourhoods V of (a,b) in U, W of $a \in \mathbb{R}^m$ and a C^1 map $g : W \to \mathbb{R}^n$ such that

$$\{(x, g(x)) \mid x \in W\} = X = \{(x,y) \in V \mid f(x,y) = 0\}$$

(i.e. locally we can solve $f(x,y) = 0$ by $y = g(x)$)

<u>Proof</u>

Consider $F(x,y) = \begin{pmatrix} x \\ f(x,y) \end{pmatrix} \in \mathbb{R}^{m+n}$, so $(x,y) \in V$, $F : V \to \mathbb{R}^{m+n}$

$DF = \begin{array}{c} \left(\begin{array}{c|c} I_m & 0 \\ \underbrace{D_1 f}_{n} & \underbrace{D_2 f}_{n} \end{array}\right) \begin{array}{l} \}m \\ \}n \end{array} \end{array}$ so $\det DF = \det D_2 f$, hence DF is invertible at (a,b) and is continuous on U. So applying the Inverse Function Theorem to F, $\exists V$ with $(a,b) \in V \subset U$ such that $F|_V : V \to F(V)$ is bijective, with C^1 inverse $G = G(x,z)$

$G(x,z) = (G_1(x,z), G_2(x,z))$. Now $F \circ G = id_{F(V)}$

so $(F \circ G)(x,z) = (G_1(x,z), f(G_1(x,z), G_2(x,z))) = (x,z)$

so $G_1(x,z) = x$.

Then $f(x, G_2(x,z)) = z$, $\forall (x,z) \in F(V)$ (1)

Also $G \circ F = id_V \Rightarrow G_2(x, f(x,y)) = y \; \forall (x,y) \in V$

Let $p_1 : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ be $p_1(x,y) = x$ and $W = p_1(V \cap f^{-1}(0))$

$W = \{x \in \mathbb{R}^m \mid \exists y \in \mathbb{R}^n \text{ with } (x,y) \in V, f(x,y) = 0\}$

(We are trying to show that $p_1 : V \cap f^{-1}(0) \xrightarrow{\sim} W$)

$g(x) = G_2(x, 0)$. Then (1) $\Rightarrow \forall x \in W$, $f(x, g(x)) = 0$.

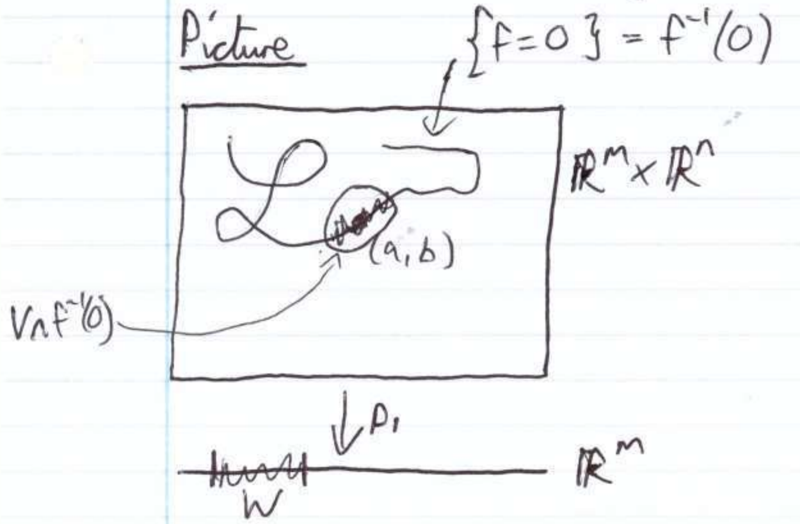
and if $(x,y) \in V$ with $f(x,y) = 0$ then $2 \Rightarrow g(x) = y$.

So this g has the required properties. $\square$

28/11/11

Analysis II (23)

Picture



$D_2 f(a,b)$ invertible

(i.e. no vertical line is tangent to $\{f=0\}$ at (a,b))

Further Remarks

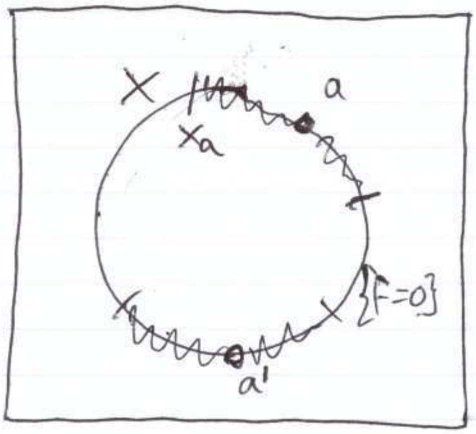
① $\mathbb{R}^{m+n} \supset U \xrightarrow{f} \mathbb{R}^n$. Suppose $a \in \mathbb{R}^{m+n}$ and rank $Df(a) = n$

Then some n columns of the matrix of $Df(a)$ are linearly independent

So the Theorem says we can use the remaining n variables "as parameters in a neighbourhood of a in $\{f=0\}$"

② As in ①, but suppose rank $Df(a) = n \quad \forall a \in X = \{f=0\}$

Regard X as a metric space (with induced metric). Then the proof shows that each $a \in X$ has an open neighbourhood $X_a \overset{\text{open}}{\subset} X$ together with $g_a : \cancel{X_a} \cong g_a(X_a) \subset \mathbb{R}^m$



$X = \bigcup X_a$

each $X_a \underset{g_a}{\overset{\cong}{\rightarrow}} g_a(X_a) \underset{\text{open}}{\subset} \mathbb{R}^m$

X is obtained by gluing various open subsets of $\mathbb{R}^m$

i.e. X is a <u>manifold</u>

30/11/11 Analysis II (24)

(B) Ordinary Differential Equations

~~TT~~ (Typical Theorem, illustrates the idea but is seldom useful)

Theorem 39 Let $F : [a,b] \times \mathbb{R}^n \to \mathbb{R}^n$ be continuous such that $\exists k$ with

$\|F(t,x) - F(t,y)\| < k\|x-y\|$ (*), $\forall t \in [a,b]$, $\forall x, y \in \mathbb{R}^n$

Let $t_0 \in [a,b]$, $x_0 \in \mathbb{R}^n$. Then $\exists$ unique conditions $f : [a,b] \to \mathbb{R}^n$

such that (1) $f(t_0) = x_0$, $\forall t \in (a,b)$, $\frac{df}{dt} = F(t, f(t))$

Proof

Transform the ODE into an integral equation (of Volterra type)

Consider $C[a,b]^n = \{\text{continuous } f : [a,b] \to \mathbb{R}^n\}$ with uniform norm

$\|f\| = \sup_{[a,b]} \|f(t)\|$ and the equation $f(t) = x_0 + \int_{t_0}^{t} F(s, f(s))\,ds$ (2)

If $f \in C[a,b]^n$, then $F(s, f(s))$ is continuous, and the Fundamental Theorem of Calculus implies that:

f differentiable and satisfies (1) $\Leftrightarrow$ f satisfies (2) $\quad \forall t \in [a,b]$

Consider $T : C[a,b]^n \to C[a,b]^n$

$T(f) = x_0 + \int_{t_0}^{t} F(s, f(s))\,ds \qquad (t \in [a,b])$

(an integral operator), so $T(f) = f \Leftrightarrow f$ satisfies (2)

Show that T is a contraction (under suitable hypotheses)

$f, g \in C[a,b]^n$, $\|T(f) - T(g)\| = \sup_{t \in [a,b]} \left\| \int_{t_0}^{t} F(s, f(s)) - F(s, g(s))\,ds \right\|$

$\|T(f) - T(g)\| \leq (b-a) \sup_{s \in [a,b]} \|F(s, f(s)) - F(s, g(s))\|$

$\leq (b-a) \sup_{s \in [a,b]} k\|f(s) - g(s)\| = (b-a)k\|f - g\|$

So$^{(a)}$ if $b - a < \frac{1}{k}$, T is a contraction. Since $C[a,b]^n$ is complete, the Contraction Mapping Theorem implies that there is a unique fixed point f i.e. a unique solution to (1).

(b) In general, decompose $a = a_0 < a_1 < \dots < a_m = b$ with $a_{i+1} - a_i < \frac{1}{k}$; use (a) to solve equation (2) on $[a_i, a_{i+1}]$ and glue them together in the obvious way (since the condition $*$ holds on the subintervals, $t_0 \in [a_j, a_{j+1}] \Rightarrow f_j$ on $[a_j, a_{j+1}]$. Use $t_0 = a_{i+1}$, $x_0 = f_i(a_{i+1})$ for $[a_{i+1}, a_{i+2}]$

Example

$\frac{df}{dt} = t^2 f + e^t \sin f$ satisfies the Lipschitz Condition $(*)$

but $\frac{df}{dt} = f^2 + 1$ does not

Theorem 40 (Picard-Lindelöf) Let $F : [a,b] \times \Omega \to \mathbb{R}^n$ be continuous where $\Omega = \bar{B}(x_0, R)$ for some $x_0 \in \mathbb{R}^n$, $R > 0$. Suppose for some $k > 0$, $\forall t \in [a,b]$, $\forall x, y \in \Omega$,

$\|F(t,x) - F(t,y)\| \leq k\|x - y\|$

Let $t_0 \in (a,b)$. Then

(a) For some $\varepsilon > 0$, $\frac{df}{dt} = F(t, f(t))$, $t \in (a,b)$, $f(t_0) = x_0$ $(**)$ has a unique continuous solution f on $[t_0 - \varepsilon, t_0 + \varepsilon] \cap (a,b)$

(b) If $\sup_{[a,b]\times\Omega} \|F\| < \frac{R}{b-a}$ then $\exists!$ continuous $f : [a,b] \to \Omega$ satisfying $(**)$ on (a,b)

N.B. Because Ω is closed and bounded, k will always exist if $F \in C'$.

Proof (b)

Let $X = \{$ continuous $f : [a,b] \to \Omega$ with uniform metric (complete by proposition 30) and $T : \cancel{C[a,b]^n} \to C[a,b]^n$,

$Tf = x_0 + \int_{t_0}^{t} F(s, f(s))\, ds$

Now we must show that $Tf \in X$. Let $m = \sup_{[a,b]\times\Omega} \|F\|$

30/4/11

Then if $f \in X$, $\| Tf(t) - x_0 \| = \| \int_0^t F(s, f(s))\, ds \| \leq (b-a) M \leq R$ by hypothesis. So $Tf \in X$. Just as in the proof of Theorem 39, $\forall f, g \in X$

$\| T(f) - T(g) \| \leq (b-a) k \| f - g \|$ (4)

So if $b - a < \frac{1}{k}$, T is a contraction, and its fixed point $f \in X$ is the unique solution to $(**)$. To get (a), just shrink $[a, b]$ until $\sup \| F \| \leq \frac{R}{b-a}$

If $(b-a) k \geq 1$ we need to modify the argument.

There are two ways: a) Example 8, Sheet 4

b) Consider the n^{th} iterate $T^n = \underbrace{T \circ \dots \circ T}_{n}$. Claim:

$\| T^n f - T^n g \| \leq \frac{(b-a)^n k^n}{n!} \| f - g \|$

As $\frac{y^n}{n!} \to 0$ as $n \to \infty$, so for sufficiently large n, T^n is a contraction, so T has a unique fixed point by Theorem 38.

We prove this claim by induction on n. We know this already for $n = 1$. $\| T^n f - T^n g \| \leq \int_{t_0}^t \| F(s, T^{n-1} f(s)) - F(s, T^{n-1} g(s)) \| ds$

$$\leq \int_{t_0}^t k \| T^{n-1} f(s) - T^{n-1} g(s) \| ds$$

$$\leq \int_{t_0}^t k \frac{(s - t_0)^{n-1}}{(n-1)!} k^{n-1} ds \text{ by the induction hypothesis (if } t > t_0)$$

$$= \frac{k^n (t - t_0)^n}{n!} \leq k^n (b-a)^n / n! \text{ (if } t < t_0 \text{ reverse the sign)}$$

<u>Example</u> $\frac{df}{dt} = f^{\frac{1}{2}}$, $t \in [0, b]$, $f(0) = 0$

$F(t, x) = x^{\frac{1}{2}}$, <u>NOT</u> Lipschitz at $x = 0$

$\forall \alpha \in [0, b]$, $f(t) = \begin{cases} 0 & 0 \leq t \leq \alpha \\ \frac{1}{4}(t - \alpha)^2 & \alpha \leq t \leq b \end{cases}$ is a solution.

So the conditions are necessary for uniqueness.

There is however an <u>existence only</u> Theorem for equations
without the Lipschitz condition (Peano)