Book has nice example illustrating issues with Bolzano-Weierstrass

Some sequences are easy to analyze. e.g. $\{\frac{1}{n}\}_{n=1}^{\infty}$ or even

$\{1, 1, 1/2, 1, 1/3, 1, 1/4, 1, \ldots\}$ but $\sin(x_n)$ for any sequence of real #'s x_n is an ∞ sequence in $[-1,1]$ so has conv. subsequence by Bolzano-Weierstrass thm.

However, if we try to find which cuts contain ∞-ly many pts, find it is very hard to do.

e.g. $x_n = 2 \cdot 10^n$

then $\sin(x_n)$ pos/neg. depending on n^{th} digit of π. Hard question...

START OF MONDAY'S LECTURE

On to proof of Thm. 1.6.9.

recall that a number S is the supremum of $f: \mathcal{U} \to \mathbb{R}$ if it is the least upper bound of the values $\{f(\underline{x}) \mid \underline{x} \in \mathcal{U}\}$

Write $S = \sup_{\underline{x} \in \mathcal{U}} f$.

upper bound: a number $\underline{a}$ s.t. $\underline{a} \geq f(\underline{x}) \quad \forall \ \underline{x} \in \mathcal{U}$.

least upper bound: if $\underline{b}$ is any other upper bound, $\underline{a} \leq \underline{b}$.

there is a corresponding notion of greatest lower bound.

Property of real numbers: Every nonempty subset $X \subset \mathbb{R}$ with upper bound has a least upper bound. (Thm 0.5.3) in book

Another statement for lower bounds.

proof of Thm 1.6.9 : Do this for maximum on compact set. Minimum pf is same.

idea — show values $f(\underline{x})$, $\underline{x} \in C$ are bounded.

hence have a least upper bound $\sup_{\underline{x} \in C} f$. Then construct a sequence converging to $\sup_{\underline{x} \in C} f$, whose limit is therefore in C a closed set.

Note here :

if $\sup_{\underline{x} \in C} f$ is attained by some $\underline{x}_0 \in C$ then clearly $f(\underline{x}_0) = \sup_{\underline{x} \in C} f$ is a maximum.

To show $\{ f(\underline{x}) \mid \underline{x} \in C \}$ is bounded, suppose not bounded and derive contradiction.

If not bounded, then $\exists$ $\{ \underline{x}_j \}$ with $| f(\underline{x}_j) | > j$. (Why?) . But by Bolzano-Weierstrass, $\exists$ subsequence $\underline{x}_{j_i}$ converging to a point $\underline{b} \in C$

Continuity of f $\Rightarrow$ if ~~near~~ $\underline{x}$ near $\underline{b}$ then $| f(\underline{x}) | < f(\underline{b}) + \epsilon$ for any $\epsilon > 0$.

but sequence constructed so that $| f(\underline{x}_j) | > j$ Pick $\underline{x}_{j_i}$ with $j_i > f(\underline{b}) + \epsilon$ Contradiction!

So f bounded.

Now find sequence $\{ \underline{x}_i \}$ with $f(x_i) \longrightarrow \sup_{\underline{x} \in C} f$. (Pick $\underline{x}_i$ in $B_{\frac{1}{i}} (\sup_{\underline{x} \in C} f)$)

then Bolzano-Weierstrass again implies conv. subsequence.

Because we're falling behind schedule a bit, ask you to read proof of fundamental theorem of algebra.

Thm: Let $z := x + iy$: complex variable $i = \sqrt{-1}$.

Given a polynomial $p(z) = z^k + a_{k-1} z^{k-1} + \cdots + a_0$ $a_i \in \mathbb{C}$

then p has a root. (a $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$)

Comments: (1) If p has a root, then repeatedly apply theorem to show p factors completely into linear factors:

$p(z) = (z - z_0)\, p_{k-1}(z)$. Then p_{k-1} of deg. $k-1$ has root, etc.

(2) Might seem remarkable that adding just $i = \sqrt{-1}$ to $\mathbb{R}$'s allows us to factor all polynomials. After all, $\sqrt{-1}$ is just a solution to the simplest example of polynomial with no roots in $\mathbb{R}$: $x^2 + 1$. But right way to think of this:

In $\mathbb{C}$, have complex unit circle $|z| = 1$. This includes points $e^{2\pi i/n}$
$= \cos \frac{2\pi}{n} + i \sin 2\pi/n$

which are roots of $z^n - 1 = 0$.

(3) (Corollary 1.6.15 of H-H)
For real polynomials, always be factored into linear polynomials (from real roots) and quadratic polynomials (from pairs of cx. conjugate roots. $\overline{z} = \overline{x+iy} := x - iy$.)

derivatives: one variable functions $f: \mathcal{U} \to \mathbb{R}$ $\quad \mathcal{U} \subseteq \mathbb{R}$ open

then for $a \in \mathcal{U}$, $\quad f'(a) = \lim_{h \to 0} \dfrac{f(a+h) - f(a)}{h}$

Two approaches to $f: \mathbb{R}^n \to \mathbb{R}^m$ or more generally $\mathcal{U} \to \mathbb{R}^m$, $\mathcal{U}$ open in $\mathbb{R}^n$

① try to generalize definition above

② use one-variable definition to understand change in f in one direction at a time. (partial derivatives!)

E.g. $\quad f: \mathbb{R}^n \to \mathbb{R}^m$

What is wrong with $\quad f'(\underline{a}) = \lim_{\underline{h} \to \underline{0}} \dfrac{f(\underline{a}+\underline{h}) - f(\underline{a})}{\underline{h}}$?

$\underline{h} \in \mathbb{R}^n$ (domain of f), $\quad$ can't compare them!

$f(\underline{a}), f(\underline{a}+\underline{h}) \in \mathbb{R}^m$ (image/range of f)

— with $|\underline{h}|$, a scalar, so that RHS is defined.

Or work ~~component by component in $\mathbb{R}^m$~~ :

$$\cancel{f: \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} f^{(1)}(\underline{x}) \\ \vdots \\ f^{(m)}(\underline{x}) \end{bmatrix}}$$

— even $|\underline{h}|$ is a problem since no longer sensitive to the path we approach along.

In particular, no longer accurate for $f: \mathbb{R}^1 \to \mathbb{R}^1$

What do we want derivative to be anyway? number? vector? matrix representing linear transformation?

Geometric intuition: best fit line to curve versus best fit ___ to surface.

derivative records slope — a 1×1 linear transformation

plane — viewed as 2×2 linear transformation (after changing coordinates)

Brilliant idea in one-variable: $f: U \to \mathbb{R}$, $U \subseteq \mathbb{R}$.

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

subtracting equation for tangent line.

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

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

note when we put abs. value here, we don't change whether limit is 0. (after all $-0 = 0$)

We say that $f: U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ is differentiable at $\underline{a} \in U$ if $\exists$ linear transformation (suggestive notation $Df(\underline{a}): \mathbb{R}^n \to \mathbb{R}^m$) such that

$$\lim_{\underline{h} \to 0} \frac{f(\underline{a}+\underline{h}) - f(\underline{a}) - \overbrace{Df(\underline{a}) \cdot \underline{h}}^{\in \mathbb{R}^m}}{|\underline{h}|} = \underline{0}$$

a vector in $\mathbb{R}^m$.

$|\underline{h}|$ a scalar in $\mathbb{R}$

idea: $Df(\underline{a})$ is best linear approximation to $f(\underline{a}+\underline{h}) - f(\underline{a})$ at $\underline{x} = \underline{a}$.

Now all linear transformations are matrices, so can we give an explicit description of $Df(\underline{a})$? Yes! Use partial derivatives...