MATH411 Notes

Nyx

November 18, 2024

0   Intro

- Prof. Machedon

- Last units on integration won't be on the final exam. Rest info are pretty much all on canvas.

1 $\mathbb{R}^n$

1.1   Re-intro to $\mathbb{R}^n$

Definition: The vector
$$u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$$
IF this comes with the following properties; addition of vectors, multiplication by a scalar, then you form a **vector space**.

Consequently, $\mathbb{R}^n$ is a vector space.

Additionally, $\mathbb{R}^n$ comes with a natural **inner product**, or scalar product, defined as

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

This satisfies the **axioms of a real inner product space**:

- *Symmetry:* $\langle u, v \rangle = \langle v, u \rangle$

- *Linearity:* $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$

- *Positive Definite:* $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0 \iff u = 0$

    - This is also what we mean by the **norm** $\|u\|$ of a vector; $\|u\|^2 = \langle u, u \rangle$
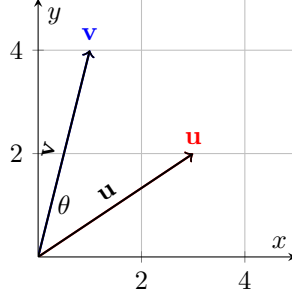
We can then define the distance of a vector:

Definition: The distance of a vector $dist(u, v)$ is defined as

$$dist(u, v) = \|u - v\|$$

Proposition: In $\mathbb{R}^2$, the inner product can be also seen as

$$\langle u, v \rangle = \|u\| \|v\| \cos(\theta)$$

where θ is the angle between the two vectors.

Proof: Using polar coordinates, $u = \|u\|(\cos(\alpha), \sin(\alpha))$ and $v = \|v\|(\cos(\beta), \sin(\beta))$. Then

$$\langle u, v \rangle = \|u\|\|v\|(\cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)) = \|u\|\|v\|\cos(\alpha - \beta) = \|u\|\|v\|\cos(\theta)$$

Definition: $u \perp v$ if $\langle u, v \rangle = 0$

Proposition: $\|u + v\|^2 = \|u\|^2 + \|v\|^2 \iff u \perp v$, eg $\langle u, v \rangle = 0$

Proof: From the definition,

$$\|u + v\|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + \langle v, v \rangle + 2\langle u, v \rangle = \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle$$

which only matches iff. $\langle u, v \rangle = 0$.

1.2 Cauchy Schwarz

Theorem (Cauchy Schwarz Inequality): Let $u, v \in \mathbb{R}^n$. Then

$$|\langle u, v \rangle| \leq \|u\|\|v\|$$

<u>Remark:</u> The proof works for any inner product satisfying the axioms, such as $\langle f, g \rangle = \int_a^b f(x)g(x)dx$. (This is a HW Problem!)

Proof: Consider that if $\|u\| = 0$ or $\|v\| = 0$, then the inequality is trivially true. Thus, suppose that $\|u\| \neq 0$ and $\|v\| \neq 0$. Then, consider

$$\|u - \alpha v\|^2$$

The goal of this proof will be to minimize α, and we will see that the Cauchy Schwarz Inequality just kind of drops out of it. We expand:

$$0 \leq \|u - \alpha v\|^2 = \|u\|^2 - 2\alpha\langle u, v \rangle + \alpha^2\|v\|^2$$

Minimizing $-2\langle u, v \rangle + 2\alpha\|v\|^2 = 0$ gives us

$$\alpha_0 = \frac{\langle u, v \rangle}{\|v\|^2}$$

Substituting this back in, we get

$$0 \leq \|u\|^2 - 2\frac{\langle u, v \rangle^2}{\|v\|^2} + \frac{\langle u, v \rangle^2}{\|v\|^4}\|v\|^2$$
$$= \|u\|^2 - \frac{\langle u, v \rangle^2}{\|v\|^2}$$
$$= \frac{\|u\|^2\|v\|^2 - \langle u, v \rangle^2}{\|v\|^2}$$

Thus, we have (by way of knowing that the denominator is ≥ 0, and so the expression as a whole being ≥ 0 requires the numerator to be as well)

$$0 \leq \|u\|^2\|v\|^2 - \langle u, v \rangle^2 \implies \langle u, v \rangle^2 \leq \|u\|^2\|v\|^2 \implies |\langle u, v \rangle| \leq \|u\|\|v\|$$

$$\square$$

<u>Remark:</u> Equality in Cauchy Schwarz occurs iff. $u = \alpha v$ for some $\alpha \in \mathbb{R}$.

One intuition is because the starting inequality we use in our proof is strict (by def of norm) unless $\|u - \alpha v\| = 0$, which only occurs when $u - \alpha v = 0$. Another way of thinking about it: Consider $u = \alpha v$. Then

$$|\langle u, v \rangle| = |\alpha|\|v\|^2 = \|\alpha v\|\|v\| = \|u\|\|v\|$$

Conversely, if $\|u\|\|v\| = \langle u, v \rangle$, then the claim is that a vector is a scalar multiple of the other. We can see this by considering the proof of the Cauchy Schwarz Inequality, and see that the entire sequence of equalities on substitution works if and only if $\langle u, v \rangle^2 = 0$ (on the last step). And thus going backwards, the norm of the vector $u - \alpha v$ is 0, which only occurs when $u = \alpha v$.

Another way of thinking about it:

$$0 \leq \|u - \alpha v\|^2 = \|u\|^2 - 2\alpha\langle u, v \rangle + \alpha^2\|v\|^2$$

This is a quadratic in α, and the only way for this to be 0 is if the discriminant is 0, which is the case iff. $u = \alpha v$.

Prof. Machedon also mentioned the method to get this by noting $\|u - \alpha_0 v\| = 0$ for $\alpha_0 = \frac{\langle u, v \rangle}{\|v\|^2} \implies u = \alpha_0 v$ by the second part of the 3rd axiom of vector spaces.

Proposition (Triangle Inequality): Let $u, v \in \mathbb{R}^n$. Then

$$\|u + v\| \leq \|u\| + \|v\|$$

Proof: Since both values are non negative, it suffices to prove that the square of the left side is less than or equal to the square of the right. Squaring the left hand side,

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v \rangle \\
&= \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle \\
&\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

The $2\langle u, v \rangle$ term is less than or equal to $2\|u\|\|v\|$ by the Cauchy Schwarz Inequality. $\square$

Proposition (Reverse Triangle Inequality): Let $u, v \in \mathbb{R}^n$. Then

$$\|u \pm v\| \geq |\|u\| - \|v\||$$

Proof: First, note that the case with $-$ implies the case with $+$ & vice versa by just replacing v with $-v$ (note that this doesn't change the norm operation on the right). Thus, we only need to prove the case with $-$.

Then we have

$$\|u\| = \|u - v + v\| \qquad\qquad\qquad \leq \|u - v\| + \|v\| \text{(Triangle Inequality)}$$

Rearranging, we get

$$\|u\| - \|v\| \leq \|u - v\|$$

$$\square$$

3

2 Limits

Let u be a limit point in $\mathbb{R}^n$. Then let $\{u_k\}_{k \in \mathbb{N}}$ be a sequence in $\mathbb{R}^n$ such that each u_k is in $\mathbb{R}^n$.

Just for this lecture, let the index be the superscript. For instance, $u_1 = (u_1^1, u_1^2, \ldots, u_1^n)$, $u_2 = (u_2^1, u_2^2, \ldots, u_2^n), \ldots$

Thus $u^i = p^i u$ as notation. TODO: clarify this point.

2.1 Limit Definitions

Definition: If $\{u_k\}$ is a sequence in $\mathbb{R}^n$, and $u \in \mathbb{R}^n$, then

$$\lim_{k \to \infty} u_k = u$$

if $\lim_{k \to \infty} \|u_k - u\| = 0$. i.e.

$$\forall \epsilon > 0, \exists K \text{ s.t. } \|u_k - u\| < \epsilon \ \ \forall k \geq K.$$

Definition: $\{u_k\} \to u$ (converges) componentwise if each $u_k^i \to u^i$ as $k \to \infty$.

Proposition: $\{u_k\} \to u$ in norm if and only if $\{u_k\} \to u$ componentwise.

Proof: Assume first $\{u_k\} \to u$ in norm. In other words,

$$\lim_{k \to \infty} \|u_k - u\| = 0$$

Then, for each fixed $1 \leq i \leq n$,

$$|u_k^i - u^i| \leq \sqrt{|u_k^1 - u^1|^2 + |u_k^2 - u^2|^2 + \ldots + |u_k^n - u^n|^2}$$
$$= \|u_k - u\| \to 0 \text{ by assumption}$$

By the comparison principle, each $\{u_k^i\} \to u^i$ as $k \to \infty$. Conversely, if each $\{u_k^i\} \to u^i$, then

$$\|u_k - u\| = \sqrt{|u_k^1 - u^1|^2 + |u_k^2 - u^2|^2 + \ldots + |u_k^n - u^n|^2}$$

We see that each part of this sum goes to 0, and thus the sum goes to 0. (by Math 410). $\qquad \square$

Remark: This only applies to finite n dimensions. For infinite dimensions, this is not true. Let's take a look at that;

Let us look at $(L^2) = \{u^1, u^2, \ldots\}$ with

$$\|u\| = \left(\sum_{i=1}^{\infty} (u^i)^2\right)^{1/2} < \infty$$

Then take a negative $\{u_k\} \in L^2$. Then $\{u_k\} \to u$ component wise if each $u_k^i \to u^i$. But $\{u_k\} \to u$ in norm if and only if $\|u_k - u\| \to 0$. Here, $\{u_k\} \to u$ componentwise but not in norm.

Take $u_1 = (1, 0, 0, \ldots), u_2 = (0, 1, 0, \ldots), u_3 = (0, 0, 1, \ldots), \ldots$. Fix an i. What is the limit $\lim_{k \to \infty} u_k^i$? It is 0. For all i in fact! Thus, we have pointwise convergence. But the norm of u_k is 1 for all k, and thus the norm of the limit is 1. Thus, we do not have convergence in norm.

3 Point Set Topology

3.1 Ballz and Open Sets

Definition: Let $r > 0, u \in \mathbb{R}^n$. Then the open ball of radius r centered at u is

$$B_r(u) = \{v \in \mathbb{R}^n : \|v - u\| < r\}$$

Definition: Let $A \subseteq \mathbb{R}^n$ be a set. Then A is open if

$$\forall u \in A, \exists r > 0 \text{ s.t. } B_r(u) \subseteq A$$

<u>Note:</u> Reading math: r depends on u, since the $\exists$ is inside the $\forall$. If you reversed this, eg $\exists r > 0 \forall u \in A$, then r would be fixed for all u. This, by the way, is the definition of the whole space (then A is all $\mathbb{R}^n$).

Proposition: All open balls are open sets.

Proof: Let the open ball be defined as $B_r(u)$. We need to show that, for some $v \in B_r(u)$, there is some $\rho > 0$ such that $B_\rho(v) \subseteq B_r(u)$.

Let us fix such a v. Then for this v,
$$\|v - u\| < r$$
by definition. Try $\rho = r - \|v - u\|$. Now we need to check that $B_\rho(v) \subseteq B_r(u)$. Let $w \in B_\rho(v)$. Then

$$\|v - w\| < \rho$$
$$= r - \|v - u\|$$

To show that $w \in B_r(u)$, look at $\|u - w\|$. By the triangle inequality,

$$\|u - w\| \leq \|u - v\| + \|v - w\|$$
$$\|u - w\| < \|u - v\| + \rho$$
$$= \|u - v\| + r - \|v - u\|$$
$$= r$$

$\square$

3.2 Closed Sets

Definition: A set $A \subseteq \mathbb{R}^n$ is closed if the following is always true:

$$\text{If } \{u_k\} \subseteq A \text{ and } \{u_k\} \to u \in \mathbb{R}^n \text{ , then } u \in A$$

Definition: The complement of a set $A \subseteq \mathbb{R}^n$ is

$$A^c = \mathbb{R}^{n \setminus A} = \{u \in \mathbb{R}^n : u \notin A\}$$

Theorem: A is open iff. A^c is closed.

Proof: Since this is an iff. statement, we prove in both directions.

$\Leftarrow$ Assume first that A is open. To show A^c is closed, let $\{u_k\} \subseteq A^c$, and assume $\{u_k\} \to u \in \mathbb{R}^n$. If $u \in A^c$, we are done. If $u \notin A^c$, eg $u \in A$, then since A is open $\exists r > 0$ such that $B_r(u) \subseteq A$. But then no u_k can be in $B_r(u)$, since $u_k \in A^c$. Thus, $u \notin A$, and thus $u \in A^c$. Thus, A^c is closed.

$\Rightarrow$ Assume now that A^c is closed. To show that A is open, pick $u \in A$. We want there to $\exists r > 0$ such that $B_r(u) \subseteq A$. Assume the contrary. Thus $\forall r > 0, B_r(u)$ contains points in A^c. Lets call those points V. Note that V is dependent on r. Let $r = \frac{1}{k}$. Thus we have some V_k. We know that $V_k \in A^c$, and $\|V_k - u\| < \frac{1}{k}$. Thus, $V_k \to u$ as $k \to \infty$. But A^c is closed, and thus $u \in A^c$. This is a contradiction, and thus A is open. $\square$

Remark: The statement A is open iff. it is not closed is not true. For instance, $\mathbb{R}^n$ is both open and closed. An example of a set that is neither open nor closed is $[0, 1)$.

Claim: $\emptyset$ is both open and closed. Checking this, $\forall u \in \emptyset$, it's trivial that $B_r(u) \subseteq \emptyset$, since anything is vacuously true for points in the empty set. Thus it's open. If $\{u_k\} \to u$ in the empty set, then it's once again vacuously true that $u \in \emptyset$. Thus it's closed.

Other sets that are both open and closed don't exist but the proof is not simple. $\mathbb{R}^n$ is path connected (future content) $\to \mathbb{R}^n$ connected, to be discussed later.

3.3   Open Set Properties

Proposition: If $\{V_s\}_{s \in S}$ is a (possibility infinite) family of open sets, then

$$\bigcup_{s \in S} V_s \text{ is open}$$

Proof: Let $u \in \bigcup_{s \in S} V_s$. Then $u \in V_{s_0}$ for some $s \in S$. Since V_{s_0} is open, $\exists r > 0$ such that $B_r(u) \subseteq V_{s_0} \subseteq \bigcup_{s \in S} V_s$. Thus, $\bigcup_{s \in S} V_s$ is open. $\square$

Proposition: If $\{V_s\}_{s \in S}$ is a finite family of open sets, then

$$\bigcap_{s \in S} V_s \text{ is open}$$

Proof: Let $u \in \bigcap_{s \in S} V_s$. Then $u \in V_s$ for all $s \in S$. Since every V_s is open, $\exists r_s > 0$ such that $B_{r_s}(u) \subseteq V_s$. Let $r = \min\{r_s : s \in S\}$. Then $B_r(u) \subseteq V_s$ for all $s \in S$, and thus $B_r(u) \subseteq \bigcap_{s \in S} V_s$. Thus, $\bigcap_{s \in S} V_s$ is open. $\square$

Note: This only works for finite families because the intersection of infinitely many open sets is not necessarily open. For instance, take $V_k = B_{1/k}(0)$. Then $\bigcap_{k=1}^{\infty} V_k = \{0\}$, which is not open.

3.4   Closed Set Properties

Proposition: Let $\{F_s\}_{s \in S}$ be a possibly infinite family of closed sets, then

$$\bigcap_{s \in S} F_s \text{ is closed}$$

Proof: We know that

$$(\bigcap_{s \in S} F_s)^c = \bigcup_{s \in S} F_s^c \text{ is open by DeMorgan's formula}$$

We also know that each F_s^c is open, and thus the union of all of them is open. Thus, the complement of the intersection is open, and thus the intersection is closed. $\square$

Proof but done another way: Let $\{u_k\} \subseteq \bigcap_{s \in S} F_s$ and $\{u_k\} \to u \in \mathbb{R}^n$. Then $u_k \in F_s$ for all $s \in S$. Since each F_s is closed, $u \in F_s$ for all $s \in S$. Thus, $u \in \bigcap_{s \in S} F_s$. Thus, $\bigcap_{s \in S} F_s$ is closed. $\square$

Proposition: Let $\{F_s\}_{s \in S}$ be a finite family of closed sets, then

$$\bigcup_{s \in S} F_s \text{ is closed}$$

Proof: We know that

$$(\bigcup_{s \in S} F_s)^c = \bigcap_{s \in S} F_s^c \text{ is open by DeMorgan's formula}$$

We also know that each F_s^c is open, and thus the intersection of all of them is open. Thus, the complement of the union is open, and thus the union is closed. $\square$

Proof but done another way: Let $\{u_k\} \subseteq \bigcup_{s \in S} F_s$ and $\{u_k\} \to u \in \mathbb{R}^n$. Note here that $\{u_k\}$ can be spread across multiple groups. However at least one F_s contains infinitely many u_k. In other words there's some F_s and a subsequence $\{u_{i_l}\}_{l \in \mathbb{N}} \to u$ contained in that F_s. Thus, $u \in F_s$ for some $s \in S$. Thus, $u \in \bigcup_{s \in S} F_s$. Thus, $\bigcup_{s \in S} F_s$ is closed. $\square$

Definition: Let $A \subseteq \mathbb{R}^n$. Look at

$$\{u \in \mathbb{R}^n | \exists r > 0 \text{ s.t. } B_r(u) \subseteq A\} \subseteq A$$

This set is called the **interior** of A, denoted A°.

Definition: Let $A \subseteq \mathbb{R}^n$. Look at

$$\{u \in \mathbb{R}^n | \exists r > 0 \text{ s.t. } B_r(u) \subseteq A^c\} \subseteq A^c$$

This set is called the **exterior** of A, denoted A^e.

Definition: Let $A \subseteq \mathbb{R}^n$. Look at

$$\{u \in \mathbb{R}^n | \ \forall r > 0 B_r(u) \text{ meets both } A \text{ and } A^c\}$$

This set is called the **boundary** of A, denoted ∂A.

4   Continuous Functions

Definition: Let $F : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Then F is continuous at $u \in A$ if

$$\forall \{u_k\} \subseteq A \text{ for which } \{u_k\} \to u, \{F(u_k)\} \to F(u)$$

Example: Take function $f : \{1, 2\} \cup [3, \infty) \to \mathbb{R}$. It will be defined by

$$f(x) = \begin{cases} 1 & \text{if x=1} \\ 0 & \text{if x=2} \\ x & \text{if } x \in [0, \infty) \end{cases}$$

Is this set continuous at $x = 1$? Yes, because all sequences in A that converge to 1 then $\{f(u_k)\}$ also converges to 1 since all $u_k = 1$ for a sufficiently large k. Basically the idea is that since it's an isolated point, all sequences that converge to 1 must be 1 after a cetain point, and thus it would also converge to $f(1)$.

Let $\epsilon = \frac{1}{10}, \exists K$ s.t. $|u_k - 1| < \frac{1}{10} \ \forall k \geq K$, but $\{u_k\} \subset \{1, 2\} \cup [3, \infty)$, and thus $u_k = 1 \ \forall k \geq K$. Thus, $f(u_k) = 1 \ \forall k \geq K$.

Proposition: Let $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}$. If f, g are continuous at $u \in A$, then

- $\alpha f + \beta g$ is continuous at u for all $\alpha, \beta \in \mathbb{R}$.

- fg is continuous at u.

- if $g(u) \neq 0 \;\; \forall u \in A$, then $\frac{f}{g}$ is continuous at u.

Proof:

- Let $\{u_k\} \to u$. Then $\{f(u_k)\} \to f(u)$ and $\{g(u_k)\} \to g(u)$. Thus, $\{\alpha f(u_k) + \beta g(u_k)\} \to \alpha f(u) + \beta g(u)$.

- Let $\{u_k\} \to u$. Then $\{f(u_k)\} \to f(u)$ and $\{g(u_k)\} \to g(u)$. Thus, $\{f(u_k)g(u_k)\} \to f(u)g(u)$.

- Let $\{u_k\} \to u$. Then $\{f(u_k)\} \to f(u)$ and $\{g(u_k)\} \to g(u)$. Since $g(u) \neq 0$, $\{g(u_k)\} \neq 0$ for sufficiently large k. Thus, $\{\frac{f(u_k)}{g(u_k)}\} \to \frac{f(u)}{g(u)}$

$\square$

Proposition: Let F, G such that $F : A \subseteq \mathbb{R}^n \to B \subseteq \mathbb{R}^n$ and $G : B \subseteq \mathbb{R}^n \to \mathbb{R}^k$. Then if F is continuous at $u \in A$ and G is continuous at $F(u)$, then $G \circ F$ is continuous at u.

Proof: Let $\{u_k\} \to u$. Then $\{F(u_k)\} \to F(u)$ and $\{G(F(u_k))\} \to G(F(u))$. Thus, $\{G(F(u_k))\} \to G(F(u))$. $\square$

Proposition: Let $F : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Then F is continuous at $u \in A$ iff. $\quad \forall \epsilon > 0 \exists \delta > 0$ such that

$$\|u - v\| < \delta \implies \|F(u) - F(v)\| < \epsilon$$

Proof: Assume $\epsilon - \delta$ definition. Let $\{u_k\} \to u$. We want that $\{F(u_k)\} \to F(u)$. Let $\epsilon > 0$. We are trying to find a K such that the distance of $F(u_k)$ to $F(u)$ is less than ϵ for all $k \geq K$. Then $\exists \delta > 0$ such that $\|u_k - u\| < \delta \implies \|F(u_k) - F(u)\| < \epsilon$. Since $\{u_k\} \to u$, $\exists K$ such that $\|u_k - u\| < \delta \;\; \forall k \geq K$.

Then, $\|F(u_k) - F(u)\| < \epsilon \;\; \forall k \geq K$. $\square$

Conversely, assume the sequence definition holds. Then, assume by contradiction that the $\epsilon - \delta$ definition does not hold. Then,

$$\exists \epsilon > 0 \text{ s.t.} \quad \forall \delta > 0, \exists v \in A, \|u - v\| < \delta \text{ s.t. } \|F(u) - F(v)\| \geq \epsilon$$

Then, take $\delta = \frac{1}{k}$, to get

$$\{v_k\} \subseteq A, \|u - v_k\| < \frac{1}{k}, \text{ and } \|F(u) - F(v_k)\| \geq \epsilon$$

Thus we found a $\{v_k\} \subseteq A$ such that $\{v_k\} \to u$ but $\{F(v_k)\} \nrightarrow F(u)$. This is a contradiction, and thus the $\epsilon - \delta$ definition holds. $\square$

<u>Recall:</u> If $F : A \to \mathbb{R}^m$, if $B \subseteq A$ then $F|_B : B \to \mathbb{R}^m$ is the restriction of F to B (also writeable as $F(B)$). Let there be a set C in $\mathbb{R}^m$. Then $F^{-1}(C) = \{u \in A : F(u) \in C\}$.

<u>Remark:</u> $\epsilon - \delta$ definition is equivalent to the idea that $\quad \forall \epsilon, \exists \delta$ such that

$$F(B_\delta(u) \cap A) \subseteq B_\epsilon(F(u))$$

intuitively this is saying that if you take a ball of the elements of A around u and map it to F, then it will be contained in a ball around $F(u)$.

This is also equivalent to

$$B_\delta(u) \cap A \subseteq F^{-1}(B_\epsilon(F(u)))$$

Theorem: Let O be open in $\mathbb{R}^n$. Then $F : O \to \mathbb{R}^m$. Then F is continuous (eg continuous at every point) on O iff.

$$F^{-1}(V) \text{ is open in } O \text{ for all open } V \subseteq \mathbb{R}^m$$

Proof: Assume F^{-1} is open for all open $V \subseteq \mathbb{R}^m$. Fix a u. We want to show that $\forall \epsilon > 0, \exists \delta > 0$ such that

$$B_\delta(u) \cap O \subseteq F^{-1}(B_\epsilon(F(u)))$$

We need to pick a V. An obvious V to pick is $B_\epsilon(F(u))$. Then $F^{-1}(B_\epsilon(F(u)))$ is open by assumption. Thus $u \in F^{-1}(B_\epsilon(F(u)))$. Thus, $\exists \delta > 0$ such that $B_\delta(u) \subseteq F^{-1}(B_\epsilon(F(u)))$. Thus, F is continuous at u. We did not use O open, but $O = F^{-1}(\mathbb{R}^m)$ is open, implied by the idea that $F^{-1}(B_\epsilon(F(u)))$ is open.

Conversely, assume F is continuous at every $u \in O$. Let $V \subseteq \mathbb{R}^m$ be open. We want to show that $F^{-1}(V)$ is open.

Let $u \in F^{-1}(V)$. We want to show that $\exists r > 0$ such that $B_r(u) \subseteq F^{-1}(V)$. We know that $F(u) \subseteq V$. Since V is open, $\exists \epsilon > 0$ such that $B_\epsilon(F(u)) \subseteq V$. Since F is continuous at u, $\exists \delta > 0$ such that $B_\delta(u) \cap O \subseteq F^{-1}(B_\epsilon(F(u))) \subseteq F^{-1}(V)$. We need to thus get rid of the O in the intersection.

Since O is open, and $u \in O, \exists \delta_1 > 0$ s.t. $B_{\delta_1}(u) \subseteq O$. Then let $r = \min(\delta, \delta_1)$. Then $B_r(u) \subseteq B_\delta(u) \cap O \subseteq F^{-1}(V)$ $\square$

Corollary: Let $f : \mathbb{R}^n \to \mathbb{R}$. If f is continous, then for a fixed $\alpha \in \mathbb{R}$, one of

$$\{u \in \mathbb{R}^n | f(u) < \alpha\}, \{u \in \mathbb{R}^n | f(u) > \alpha\}$$

is open.

<u>Remark:</u>

$$\{u \in \mathbb{R}^n | f(u) \leq \alpha\}$$

is closed because its complement,

$$\{u \in \mathbb{R}^n | f(u) > \alpha\}$$

is open. This is because f is continuous, and thus the inverse image of an open set is open.

5 Compact Sets

5.1 Sequential Compactness

Definition: The set $K \subseteq \mathbb{R}^n$ is <u>sequentially compact</u> if $\forall \{u_l\} \subseteq K, \exists \{u_{l_m}\}$ and $u \in K$ s.t. $\{u_{l_m}\} \to u$.

Definition: The set $A \subseteq \mathbb{R}^n$ is <u>bounded</u> if $\exists M$ s.t. $\|u\| \leq M$ $\forall u \in A$.

Proposition: If $K \subseteq \mathbb{R}^n$ is sequentially compact, then K is bounded.

Proof: We want to find $M > 0$ s.t. $\|u\| \leq M$ $\forall u \in K$. Assume by contradiction that K is not bounded. Then $\forall M > 0, \exists u \in K$ s.t. $\|u\| > M$.

Specify $M = m \in \mathbb{N}$. Then we find some $u_m \in K$ s.t. $\|u_m\| > m$.

Then, $\{u_m\} \subseteq K$, but there is no subsequence that converges to a point in K. This is a contradiction, and thus K is bounded. $\square$

Proposition: If $K \subseteq \mathbb{R}^n$ is sequentially compact, then K is closed.

Proof: Pick some $\{u_l\} \subseteq K$. Assume that $\{u_l\} \to u \in \mathbb{R}^n$. We want some $u \in K$. Since K is sequentially compact, $\exists \{u_{l_m}\} \to u \in K$. Thus, K is closed. $\qquad\square$

Note: I'm going to stop using italic l here because i cannot be asked to MIT macro every time.

Recall: If $\{x_l\} \subseteq \mathbb{R}$, and $\{x_l\}$ is bounded, then $\exists \{x_{l_m}\} \to x \in \mathbb{R}$

Proposition: If $\{u_l\} \subseteq \mathbb{R}^n$, and $\{u_l\}$ is bounded, then $\exists \{u_{l_m}\} \to u \in \mathbb{R}^n$ (eg. every bounded sequence has a convergent subsequence).

Proof: We will use induction.

- $n = 1$, this is clearly true lol

- Assume the IH. Then $\exists \{u_l\} \subseteq \mathbb{R}^n$ that has a convergent subsequence.

- Let $\{u_l\} \subseteq \mathbb{R}^{n+1}$. Also, $\{u_l\}$ is bounded. We write that

$$u_l = (u_l^1, u_l^2)$$

  where u_l^1 is of ""dimension"" n and u_l^2 is of dimension 1.

- By the IH $\exists \{u'_{l_m}\} \to u' \in \mathbb{R}^n$.

- Look at $\{u_{l_m}^{n+1}\}$, which is a bounded sequence. Then $\exists \{u_{l_{m_p}}^{n+1}\} \to (u', u^{n+1}) \in \mathbb{R}$.

Note: What the fuck, I'm gonna have to look back at this or find it in the textbook lol

Theorem: $K \subseteq \mathbb{R}^n$ is sequentially compact iff. K is closed and bounded.

Proof: We have already shown that if K is sequentially compact then it's closed and bounded. Let K be closed and bounded.

Pick some $\{u_l\} \subseteq K$. Then $\{u_k\}$ is bounded since it lives inside a bounded set, and thus $\exists \{u_{l_m}\} \to u \in \mathbb{R}^n$. Since K is closed, $u \in K$ and we are done. $\qquad\square$

Proposition: If K is sequentially compact and $F : K \to \mathbb{R}^n$ is continuous then $F(K)$ is sequentially compact.

Proof: Let $\{u_l\} \subseteq F(K)$. Then $\exists \{u_k\} \subseteq K$ s.t. $F(u_k) = v_k$. Then $\exists \{u_{k_l}\} \to u \in K$, and $\{F(u_{l_k})\} \to F(u)$ by F being continuous. $\qquad\square$

Question: If $C \subseteq \mathbb{R}^n$ is closed, $F : C \to \mathbb{R}^m$ is continuous, is $F(C)$ closed? No. Take for instance $\frac{1}{1+x}$.

Question: If $B \subseteq \mathbb{R}$ is bounded, and $f : B \to \mathbb{R}$ is continuous, is $f(B)$ bounded? No. Take for instance $f(x) = \frac{1}{x}$, $B = (0, 1)$. Ten $f(B) = (1, \infty)$ which is not bounded.

Question: If $B \subseteq \mathbb{R}$ is bounded and $f : \mathbb{R} \to \mathbb{R}$ is continuous, is $f(B)$ bounded? Yes, take a compact set $B \subseteq K$. K can just be $[a, b]$ where a, b are the bounds of B. Then K is seq compact and $f(K)$ is bounded.

Theorem: Let K be sequentially compact. Then $f : K \to \mathbb{R}$ is continuous. Then f attains its minimum and maximum on K.

Proof: We know that $f(K)$ is bounded, and that $\exists m = \inf_{u \in K} f(u)$ (infimum). We want to show that $\exists u \in K$ s.t. $f(u) = m$. We know that $\exists \{u_i\} \subseteq K, f(u_i) \to m$. Since K is sequentially compact, $\exists \{u_{i_k}\} \to u \in K$. Since f is continuous, $f(u_{i_k}) \to f(u)$. Since $f(u_{i_k}) \to m$, $f(u) = m$. $\qquad\square$

Theorem: Let $A \subseteq \mathbb{R}^n$ be such that any $f : A \to \mathbb{R}$ attaining its minimum and maximum values. Then A is sequentially compact.

Proof: We want to show that A Is bounded and closed.

- Bounded - Choose $f(x) = \|x\|$. We know that $\exists x_{\max} \in A$ s.t. $f(x) \leq f(x_{\max}) \quad \forall x \in A$.

  i.e. $0 \leq \|x\| \leq \|x_{\max}\| \quad \forall x \in A$. Thus, A is bounded.

- Closed - Let $\{u_i\} \in A, \{u_i\} \to u \in \mathbb{R}^n$. We want $u \in A$. Let $f(x) = \|x - u\|$. We choose this function because we wanted a function which has 0 as an infimum. We also see that $f(u_i) \to 0$. Thus, $\exists u_i$ s.t. $f(u_i) \leq f(u_i) \leq f(u_{\max})$. Thus, $u \in A$.

5.2 Uniform Continuity

Definition: $F : A \to \mathbb{R}^m$ is <u>uniformly continuous</u> if $\quad \forall \{u_i\}, \{v_i\} \subseteq A$ if $\{\|u_i - v_i\|\} \to 0$, then $\{\|F(u_i) - F(v_i)\|\} \to 0$.

<u>Remark:</u> F uniformly continuous implies F continuous. The converse is not true. For instance, $f(x) = x^2$ is continuous but not uniformly continuous. In general, not all continuous functions are uniformly continuous.

x^2 is continuous but not uniformly continuous. Let $u_i = i, v_i = \frac{1}{i} + i$. Then $\|u_i - v_i\| \to 0$ but $\|f(u_i) - f(v_i)\| \nrightarrow 0$.

Not all continuous $f : (0, 1) \to \mathbb{R}$ are uniformly continuous. Take for instance $u_i = \frac{1}{i}, v_i = \frac{1}{i^2}$. Then $f = \frac{1}{x}$ is continuous but not uniformly continuous.

<u>Recall:</u> $F : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is uniformly continuous if $\quad \forall \{u_i\}, \{v_i\} \subseteq A$, with $\|u_i - v_i\| \to 0$, then $\|F(u_i) - F(v_i)\| \to 0$.

Theorem: Let $K \subseteq \mathbb{R}^n$ be sequentially compact, and $F : K \to \mathbb{R}^m$ be continuous. Then $F : K \to \mathbb{R}^m$ is uniformly continuous.

Proof: Let $F : K \to \mathbb{R}^m$ continuous. Then assume by contradiction that F is not uniformly continuous, i.e. $\exists \{u_i\}, \{v_i\} \subseteq K$ such that $\|u_i - v_i\| \to 0$ but $\|F(u_i) - F(v_i)\| \nrightarrow 0$.

This means that $\quad \forall \epsilon > 0, \exists I$ s.t. $\|F(u_i) - F(v_i)\| \geq \epsilon \quad \forall i \geq I$. We find a subsequence $\{u_{i_j}\}$ with $\|F(u_{i_j}) - F(v_{i_j})\| \geq \epsilon$. Since K is sequentially compact, $\exists \{u_{i_{j_k}}\} \to u \in K$. Since F is continuous, $F(u_{i_{j_k}}) \to F(u)$. But this is a contradiction because $\|F(u_{i_{j_k}}) - F(v_{i_{j_k}})\| \geq \epsilon$ for all k. $\qquad \square$

Theorem: Let $F : A \to \mathbb{R}^m$. Then the following are requivalent:

1. if $\{u_i\}, \{v_i\} \subseteq A$ and $\{\|u_i - v_i\|\} \to 0$, then $\{\|F(u_i) - F(v_i)\|\} \to 0$.

2. $\quad \forall \epsilon > 0 \exists \delta > 0$ s.t. $\|u - v\| < \delta \implies \|F(u) - F(v)\| < \epsilon$

Proof: $2 \to 1$: Assume 2. Let $\{u_i\}, \{v_i\} \subseteq A, \|u_i - v_i\| \to 0$. We want to show that $\|F(u_i) - F(v_i)\| \to 0$. Let $\epsilon > 0$. Then $\exists I$ s.t. $\|u_i - v_i\| < \delta \quad \forall i \geq I$. Then $\|F(u_i) - F(v_i)\| < \epsilon \quad \forall i \geq I$.

$1 \to 2$: Assume 1. Equivalently, if 2 is not true, then 1 cannot be true. Thus, assume 2 is not true. Then $\exists \epsilon > 0$ s.t. $\quad \forall \delta > 0, \exists u, v \in A$ s.t. $\|u - v\| < \delta$ but $\|F(u) - F(v)\| \geq \epsilon$.

Let $\delta = \frac{1}{k}$. Then $\exists u_k, v_k \in A$ s.t. $\|u_k - v_k\| < \frac{1}{k}$ but $\|F(u_k) - F(v_k)\| \geq \epsilon$. Then $\{u_k\}, \{v_k\} \subseteq A$ and $\{\|u_k - v_k\|\} \to 0$ but $\{\|F(u_k) - F(v_k)\|\} \nrightarrow 0$. $\qquad \square$

5.3 Path Connected/Intervals

Definition: A is <u>convex</u> if $\quad \forall u, v \in A$, the line segment $tu + (1-t)u$ between u and v is in $A \quad \forall 0 \le t \le 1$.

Definition: $A \subseteq \mathbb{R}^n$ is <u>path connected</u> if $\quad \forall u, v \in A$, there exists a continuous function $\gamma : [a, b] \to A$ such that $\gamma(a) = u, \gamma(b) = v$. Then γ is called the parametrized path.

Theorem: If $A \subseteq \mathbb{R}^n$ is path connected and $F : A \to \mathbb{R}^m$ is continuous, then $F(A)$ is path connected.

Proof: Let $u, v \in F(A)$. We want to find a parametrized path from u to v. Let $x, y \in A$ such that $F(x) = u, F(y) = v$. Since A is path connected, $\exists \gamma : [a, b] \to A$ such that $\gamma(a) = x, \gamma(b) = y$. Then $F \circ \gamma : [a, b] \to F(A)$ is a parametrized path from u to v. $\qquad \square$

5.4 Connectedness

Proposition: $A \subseteq \mathbb{R}$ is path connected iff. A is an interval I.

Proof: Assume A is an interval. Let $u, v \in A$. Then let $\gamma : [0, 1] \to A$. Then $\gamma(t) = tu + (1-t)v$ is a parametrized path from u to v. Thus A is path connected.

Assume A is path connected. Let $u, v \in A$. We will show that $[u, v] \in A$. We know that $\exists \gamma : [a, b] \to A$ such that $\gamma(a) = u, \gamma(b) = v$. Let $w \in [u, v]$. We want that $w \in A$. By the IVT, $\exists c \in [a, b]$ s.t. $\gamma(c) = w$. Thus, $w \in A$ (since $\gamma : [a, b] \to A$). $\qquad \square$

Definition: A has the intermediate value property (IVP) if $\quad \forall f : A \to \mathbb{R}$, continuous, then $f(A)$ is an interval.

Definition $A \subseteq \mathbb{R}^n$ is <u>not connected</u> if $\exists U, V$ open sets such that

1. $U \cap A \ne \emptyset, V \cap A \ne \emptyset$

2. $(U \cap A) \cap (V \cap A) = \emptyset$

3. $(U \cap A) \cup (V \cap A) = A$

<u>Remark:</u> $U \cap A$ is called relatively open. If U, V satisfy 1,2,3, then U, V are said to separate A.

Theorem: A is connected iff. A has the IVP ($\forall f : A \to \mathbb{R}$ continuous, then $f(A)$ is an interval).

Proof: We will show that A is not connected is equivalent to A does not have the IVP ($\exists f : A \to \mathbb{R}$, continuous, such that $f(A)$ is not an interval).

Assume that A is not connected. Then let U, V open separate A. Let

$$f(x) = \begin{cases} 1 & \text{if } x \in A \cap U \\ 0 & \text{if } x \in A \cap V \end{cases}$$

f is defined $\quad \forall x \in A$ because of 3. The value $f(x)$ is unique because of 2. $f(A) = \{0, 1\}$ because of 1. To show that f is continuous at some $x_0 \in U \cap A$, we need to show that $\exists \delta > 0$ such that $B_\delta(x_0) \subseteq U$ (which is true, since U is open). Then if $x \in A$ and $\|x - x_0\| < \delta$, then $f(x) = f(x_0)$. Then $|f(x) - f(x_0)| = 0 < \epsilon$. We have thus found $f : A \to \mathbb{R}$ continuous, $f(A)$ not an interval.

Conversely, assume A does not have the IVP. Then let $f : A \to \mathbb{R}$ be continuous such that $f(A)$ is not an interval. Then $\exists c \in \mathbb{R}$ such that $c \notin f(A), f(A) \cap (-\infty, c) \ne \emptyset, f(A) \cap (c, \infty) \ne \emptyset$.

To construct U, V separating A, let

$$\tilde{U} = f^{-1}((-\infty, c)), \tilde{V} = f^{-1}((c, \infty))$$

Then $\tilde{U}, \tilde{V}$ are not the empty set, because of the property that $f(A) \cap (-\infty, c) \neq \emptyset$ and $f(A) \cap (c, \infty) \neq \emptyset$. Then $\tilde{U} \cup \tilde{V} = A$ because of the property that $c \notin f(A)$. To finish the proof, we will find U, V open such that

$$\tilde{U} = U \cap A, \tilde{V} = V \cap A$$

Let $x \in \tilde{U}$ Then $f(x) < c$. Since f is continuous, $\exists r(x) > 0$ such that $f(y) < c \;\; \forall y \in B_{r(x)}(x) \cap A$. Let $U = \bigcup_{x \in \tilde{U}} B_{r(x)}(x)$. Note that $\tilde{U} = U \cap A = \bigcup_{x \in \tilde{U}} (B_{r(x)}(x) \cap A)$. If $x \in \tilde{U}$, then $x \in B_{r(x)}(x) \cap A$. If $y \in B_{r(x)}(x) \cap A$, for some $x \in \tilde{U}$, then $y \in \tilde{U}$. Thus we are done $\qquad \square$

Consequences:

- A being path connected implies A has IVP is equivalent to A connected

- A connected does not imply path connected

Corollary: If $U \subseteq \mathbb{R}^n$ is both open and closed, then $U = \mathbb{R}^n$ or $U = \emptyset$.

Proof: $\mathbb{R}^n$ is path connected implies $\mathbb{R}^n$ is connected. Let U be both open and closed. Let $V = U^c$. Then V is also open and closed. Thus $U \cup V = \mathbb{R}^n$. We have that $U \cap V = \emptyset, U \cup V = \mathbb{R}^n$. Is it possible for both of them to not be the empty set? No, because $\mathbb{R}^n$ is connected. Thus $U = \emptyset$ or $U = \mathbb{R}^n$. $\qquad \square$

Definition: $K \subseteq \mathbb{R}^n$ is $\underline{\text{compact}}$ if $\;\; \forall$ family of open sets $\{V_\alpha\}$, if $K \subseteq \bigcup_\alpha V_\alpha$, then $\exists \{V_{\alpha_i}\}$ finite such that $K \subseteq \bigcup_{i=1}^n V_{\alpha_i}$.

Theorem: K compact iff. K is closed and bounded.

Proposition 1: K compact implies K bounded.

Proof: Our goal is to contain K in the union of a finite amount of sets. Let $V_\alpha = B_\alpha(0), \alpha > 0$. Then $\exists \alpha_1 < \alpha_2 < \alpha_l$ such that $K \subseteq \bigcup_{i \in \alpha} V_{\alpha_i} = B_{\alpha_n}(0)$.

Proposition 2: K compact implies K closed.

Proof: Let K compact. Then let $\{u_k\} \in K, \{u_k\} \to u \in \mathbb{R}^n$. We want to show that $u \in K$. Assume, by contradiction, that $u \notin K$. Let $V_l = \{x \in \mathbb{R}^n | \|x - u\| > \frac{1}{l}\}$. Then $\bigcup_{l=1}^\infty V_l = \mathbb{R}^n \setminus \{u\}$. Also, $K \subseteq \bigcup_{l=1}^\infty V_l, V_1 \subseteq V_2 \subseteq \ldots$. Thus $\exists V_{l_0}$ s.t. $K \subseteq V_{l_0} = \{x \in \mathbb{R}^n | \|x - u\| > \frac{1}{l_0}\}$. This is a contradiction because we don't have $\{u_k\} \to u$. Thus $u \in K$. $\qquad \square$

Proposition: If L is compact, and $K \subseteq L$, and K is closed, then K is compact.

Proof: Let V_α be open sets, with $K \subseteq \bigcup_\alpha V_\alpha$. Then $L \subseteq \bigcup_\alpha V_\alpha \cup K^c$. If $x \in K$, then $x \in \bigcup_\alpha V_\alpha$. If $x \notin K$, then $x \in K^c$. That's why we needed the K^c in the statement earlier.

Since L is compact,

$$\exists V_{\alpha_1}, \ldots, V_{\alpha_n} \text{ s.t. } K \subseteq V_{\alpha_1} \cup \cdots \cup V_{\alpha_n} \cup K^c$$

We of course want to get rid of the K^c. But by definition nothing in K is in K^c. Thus, we have that K is compact. $\qquad \square$

Theorem: If $K \subseteq \mathbb{R}^n$ is closed and bounded, then K is compact.

Proof: Since K is bounded, K is contained in some closed cube in $\mathbb{R}^n$. For simplicity, assume that $K \subseteq [0,1] \times [0,1] \cdots \times [0,1] = C$. By the previous proposition, it suffices to show taht C is compact. Let $\{V_\alpha\}$ be a family of open sets such that $C \subseteq \bigcup_\alpha V_\alpha$ eg. it is an open cover of C.

Assume by contradiction there does not exist V_{α_i} finite subcover ($K \subseteq V_{\alpha_1} \cup V_{\alpha_2} \cup \cdots \cup V_{\alpha_n}$ never happens).

Divide C into 2^n cubes of the form $C_i = [\frac{j_1}{2}, \frac{j_1+1}{2}] \times [\frac{j_2}{2}, \frac{j_2+1}{2}] \cdots \times [\frac{j_n}{2}, \frac{j_n+1}{2}]$ where $j_i = 0, 1$. So then there exists a subcube WLOG C_1 that cannot be covered by a finite amount of V_α.

Continuing, we will eventually find some subcube C_2, then C_3, and so on. None of these C_i can be covered by finitely many V_α. By a generalization of the Nested Interval Theorem to $\mathbb{R}^n$, $\exists x_0 \in C$ such that $x_0 \in C_i$ for all i. Thus, $x_0 \in C$ cannot be covered by any finite amount of V_α. This is a contradiction. Thus, K is compact. $\qquad\square$

6 Metric Spaces

6.1 Blah

Definition: A set X and $d : X \times X \to \mathbb{R}$ is a <u>metric space</u> if

1. $d(u, v) = d(v, u) \quad \forall u, v \in X$ (symmetry)

2. $d(u, v) \geq 0$ and $d(u, v) = 0 \iff u = v$ (non-negativity)

3. $d(u, v) \leq d(u, w) + d(w, v)$ (triangle inequality)

<u>Example:</u> $X = \mathbb{R}^n, d(u, v) = \|u - v\|$ is a metric space.

<u>Remark:</u> In general, if V is a vector space, $\|\cdot\|$ is called a norm if

1. $\|u\| \geq 0$

2. $\|u\| = 0 \iff u = 0$

3. $\|u + v\| \leq \|u\| + \|v\|$

4. $|\alpha|\|u\| = \|\alpha u\|$

Proposition: If $\|\cdot\|$ is a norm, on a vector space V, then $d(u, v) = \|u - v\|$ is a metric.

Proof: To do so we need to prove each property.

1. $d(u, v) = \|u - v\| = \|v - u\| = d(v, u)$ by property 4 of norms.

2. $d(u, v) = \|u - v\| \geq 0$ by property 1 of norms. If $d(u, v) = 0$, then $\|u - v\| = 0 \implies u = v$ by property 2 of norms.

3. $d(u, v) = \|u - v\| \leq \|u - w\| + \|w - v\| = d(u, w) + d(w, v)$ by property 3 of norms.

6.2 Practice Problems Time

- Let $f : \mathbb{R} \to \mathbb{R}$ continuous. Prove that $G = \{(x, f(x)) | x \in \mathbb{R}\}$ is closed.

  - By definition, this means that if there's a sequence of points and it converges to something then it must be in G. Let $\{(x_n, f(x_n))\} \to (x, y)$. We want to show that $(x, y) \in G$.

    $x_n \to x$ because f is continuous, and also $f(x_n) \to f(x)$. Since f is continuous, $y_n = f(x_n) \to f(x)$ because it's in G. Thus, $(x, y) = (x, f(x)) \in G$.

- Is the converse to this true? eg. if you have a function $f : \mathbb{R} \to \mathbb{R}$ such that $G = \{(x, f(x)) | x \in \mathbb{R}\}$ is closed, does it imply that f is continuous?

- Traditional thoughts: Attempt to prove it first. We have that $\{x_k\} \to x$, we hope to show that $f(x_k) \to f(x)$. We know that $\{(x_k, f(x_k))\} \subseteq G$ is closed. Then if $\{(x_k), f(x_k)\} \to (x, y)$, then $f(x) = y$.

  We know that $\{x_k\} \to x$. If $\exists y$ s.t. $\{f(x_k)\} \to y$, then $y = f(x)$, and thus f is continuous.

  If there is a counterexample, it has the feature that $\{f(x_k)\}$ does not converge to a real number. An example of this is $f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$ The x can converge to a point, but the $f(x)$ does not converge to anything. Thus, the converse is not true.

  There's also $f(x) = \frac{1}{x}$

- How about if the graph domain was limited to $[0, 1]$?

  - The answer is that this is true, but it's hard to prove right now.

- Let $C \subseteq \mathbb{R}^n$ be a set with the property that any $f : C \to \mathbb{R}$ which is continuous has to be uniformly continuous. Then, prove or disprove: C is a bounded set. *Hint:* $f(x) = x^2$ is not uniformly continuous if $n, n + \frac{1}{n} \in C$ $\forall n$. Prove or disprove: C is closed.

  - No; Let's try $C = \mathbb{N}$. Let $f : \mathbb{N} \to \mathbb{R}$ be some arbitrary function. Then f is continuous automatically. Suppose that we have $\|x - y\| < \frac{1}{2}$. Then $|f(x) - f(y)| < \epsilon$ is true, since the only way this is possible is for $x = y$. Thus, C is not bounded.

  - Yes: Let's try to prove that C is closed. Let $\{x_k\} \subseteq C$ and $x_k \to x$. We want to show that $x \in C$. On the contrary let $x \notin C$. Try $f(x) = \frac{1}{x}$. Then f is not uniformly continous. For instance, let $\epsilon = 1$. Then does there $\exists \delta > 0$ such that $\|x - y\| < \delta \implies |f(x) - f(y)| < \epsilon$? No, let $x = x_k, y = x_l$, chosen such that $\|x_k - x_0\| < \frac{\delta}{2}, \|x_l - x_0\| < \frac{\delta}{2}$. Then $|f(x_k) - f(x_l)| \to \infty$. Thus, C is not closed.

6.3   Metrics

Definition: Let $\| \cdot \|$ be a norm on a vector space V. Then $d(u, v) = \|u - v\|$ is a metric on V. (we already went over this definition but ok)

Examples: (in $\mathbb{R}^2$)

- $\|x\| = \sqrt{x_1^2 + x_2^2}$ is a norm. Then $d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ is a metric.

  - Also, $B_1(0) = \{x \in \mathbb{R}^2 | \|x\| < 1\}$ uses this metric.

- $\|x\|_1 = |x_1| + |x_2|$ is a norm. Then $d(x, y) = \|x - y\| = |x_1 - y_1| + |x_2 - y_2|$ is a metric.

  - Let's convince ourselves this is a norm. Then

  - $\|x\| \geq 0, \|x\| = 0 \iff x = 0$

  - $\|\alpha x\| = |\alpha x_1| + |\alpha x_2| = |\alpha|(|x_1| + |x_2|) = |\alpha| \|x\|$

  - $\|x + y\| = |x_1 + y_1| + |x_2 + y_2| \leq |x_1| + |y_1| + |x_2| + |y_2| = \|x\| + \|y\|$

- $\|x\|_\infty = \max(|x_1|, |x_2|)$ is a norm. Then $d(x, y) = \|x - y\| = \max(|x_1 - y_1|, |x_2 - y_2|)$ is a metric.

  - Let's convince ourselves this is a norm. Then

  - $\|x\| \geq 0, \|x\| = 0 \iff x = 0$

$$- \quad \|\alpha x\| = \max(|\alpha x_1|, |\alpha x_2|) = |\alpha|\max(|x_1|, |x_2|) = |\alpha|\|x\|$$

$$- \quad \|x + y\| = \max(|x_1 + y_1|, |x_2 + y_2|) \leq \max(|x_1| + |y_1|, |x_2| + |y_2|) = \|x\| + \|y\|$$

Note: $\|x\|_\infty \leq \|x\|_1 \leq \|x\|$ Also, all norms are comparable. eg. They can all be bounded, a sequence converging under one norm will converge under another norm, etc.

Definition: $X = C([a,b], \mathbb{R}) = \{f : [a,b] \to \mathbb{R} | f \text{ continuous }\}$ is the set of continuous functions from $[a,b]$ to $\mathbb{R}$.

Example: In ∞ dimensions:

Let $X = C([a,b], \mathbb{R}) = \{f : [a,b] \to \mathbb{R}, f \text{ continuous }\}$ where C is the set of continuous functions.

- Then $\|f\| = \|f\|_\infty = \max_{x \in [a,b]} |f(x)|$ is a norm.

  $- \quad max_{x \in [a,b]} |f_1(x) + f_2(x)| \leq max_{x \in [a,b]} |f_1(x)| + max_{x \in [a,b]} |f_2(x)| = \|f_1\| + \|f_2\|$

- $\|f\|_1 = \int_a^b |f(x)| dx$ is a norm.

  $- \quad \int_a^b |f_1(x) + f_2(x)| dx \leq \int_a^b |f_1(x)| dx + \int_a^b |f_2(x)| dx = \|f_1\| + \|f_2\|$

  $- \quad \int_a^b |\alpha f(x)| dx = |\alpha| \int_a^b |f(x)| dx = |\alpha|\|f\|$

  $- \quad \int_a^b |f(x)| dx \geq 0, \int_a^b |f(x)| dx = 0 \iff f(x) = 0$

- $\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx}$ is a norm.

  $- \quad \sqrt{\int_a^b |f_1(x) + f_2(x)|^2 dx} \leq \sqrt{\int_a^b |f_1(x)|^2 dx} + \sqrt{\int_a^b |f_2(x)|^2 dx} = \|f_1\| + \|f_2\|$

  $- \quad \sqrt{\int_a^b |\alpha f(x)|^2 dx} = \sqrt{\alpha^2 \int_a^b |f(x)|^2 dx} = \alpha \sqrt{\int_a^b |f(x)|^2 dx} = \alpha\|f\|$

  $- \quad \sqrt{\int_a^b |f(x)|^2 dx} \geq 0, \sqrt{\int_a^b |f(x)|^2 dx} = 0 \iff f(x) = 0$

Remark: If Y is a metric space with metric d, then any subset $X \subseteq Y$ is also a metric space

Generalities:

- if X is a metric space, the ball $B_r(p) = \{q \in X | d(p,q) < r\}$

- A in X is open if $\quad \forall p \in A, \exists r > 0$ s.t. $B_r(p) \subseteq A$

Definition: $A \subseteq X$ is closed if $\quad \forall \{p_k\} \subseteq A$, if $\{p_k\} \to p \in X$, then $p \in A$.

Remark: If X is a metric space, then X is open and closed in X.

Example: $X = [0,1)$ with $d(x,y) = |x - y|$ has X open and closed in itself.

Remark: $\{f_k\} \subseteq C([a,b], \mathbb{R}), \{f_k\} \to f$ with respect to $d(f,g) = \max_{x \in [a,b]} |f(x) - g(x)|$ implies that $f_k \to f$ uniformly.

Recall from 410 that uniformly means that $\quad \forall \epsilon > 0, \exists N$ s.t. $\|f_k - f\| < \epsilon \quad \forall k \geq N$.

Definition: $\{p_k\} \subseteq X$ is <u>Cauchy</u> (or is a Cauchy sequence) if

$$\forall \epsilon > 0, \exists K \text{ s.t. } d(p_k, p_l) < \epsilon \quad \forall k, l \geq K$$

Definition: X is <u>complete</u> if

$$\forall \text{ Cauchy seq } \in X \exists p \in X \text{ s.t. } \{p_k\} \to p$$

<u>Example:</u> For instance, the uniformly remark earlier has complete.

<u>Comparison:</u> Take

$$\int_a^b |f(x)| dx$$

for $f \in C([a, b], \mathbb{R})$. Then take

$$\max_{x \in [a,b]} |f(x)|$$

There is one clear inequality here. This is that

$$\int_a^b |f(x)| dx \leq (b - a) \max_{x \in [a,b]} |f(x)|$$

Does there $\exists c > 0$ such that

$$\max_{x \in [a,b]} |f(x)| \leq c \int_a^b |f(x)| dx$$

No, take a function over 0 to 1 that drops from 1 to 0 very quickly. Then the max is 1, but the integral is very small. Thus, this inequality does not hold.

<u>Question:</u> For $C([0, 1], \mathbb{R})$ with $d(f, g) = \int_0^1 |f - g|$, is this complete?

One of the axioms of the norm is that $\|f\| = 0 \implies f = 0$. As a remark, if f is continuous, that's true. If f is a general Riemann integrable function, then the above may fail. For instance, take the function which is 0 everywhere except for one place, where it is 1. Then the integral is 0, but the function is not 0. So this is not complete.

Let $f_n(x) = x^n$, $f_n : [0, 1] \to \mathbb{R}$. Then $\{f_n\}$ will approach 0 if $x \in [0, 1)$, and 1 if $x = 1$. Then f_n goes to 0 with respect to $d_1(f_n, 0) = \int_0^1 |f_n - 0| dx$. And f_n is not cauchy with respect to $d(f, g) = \max_{x \in [0,1]} |f(x) - g(x)|$.

Then, let

$$f_n(x) = \begin{cases} x^n & x \in [0, 1] \\ 1 & x \in [1, 2] \end{cases}$$

Then

$$\int_0^2 |f_n(x) - f(x)| dx \to 0 \text{ if } f(x) = \begin{cases} 0 & x \in [0, 1) \\ 1 & x \in [1, 2] \end{cases}$$

eg $\int_0^1 |f_n(x) - 0| dx \to 0, \int_1^2 |f_n(x) - 1| dx \to 0$

17

So $f_n \to f$ with respect to d, therefore it is Cauchy. But there $\nexists g \in C$ such that $\int_0^2 |f - g| dx = 0$. Thus, this is not complete. This is why we have to give up the Riemann integral and move to the Lebesgue integral.

<u>Exercise:</u> Let $f_n \in C([0,1], \mathbb{R})$, $f_n(x) = x^n$. Then $d(f,g) = \max_{x \in [0,1]} |f(x) - g(x)|$. This is not Cauchy.

How about $x^n(1-x)$? Is this Cauchy? Eg. Does it converge uniformly to something? For x strictly less than 1, it converges to 0, and if $x = 1$, it's identically 0. Thus the candidate for a limit is $f(x) = 0 \quad \forall x \in [0,1]$.

The issue is: $\max_{x \in [0,1]} x^n(1-x) \to 0$.

<u>Exercise (12.1, #4)</u> Let $f_k \in C([0,1], \mathbb{R})$, $f_k(x) = \cos(\frac{x}{k})$. Show that this converges uniformly.

It is sufficient to show that the maximum of the absolute value of the difference between the function and the limit candidate goes to 0.

7 Existence and Uniqueness for ODEs

7.1 Uh Oh.

Definition: Let X be a metric space, and $T : X \to X$ be a map. Then T is <u>lipschitz</u> if $\exists M > 0$ s.t. $d(T(p), T(q)) \leq M \cdot d(p, q) \quad \forall p, q \in X$.

Proposition: If $f : \mathbb{R} \to \mathbb{R}$, differentiable, then f is lipschitz with constant M iff.

$$|f'(x)| \leq M \quad \forall x \in \mathbb{R}$$

Proof: If f is differentiable, and lipschitz with constant M, then

$$|f(x) - f(y)| \leq M|x - y| \quad \forall x, y \in \mathbb{R}$$
$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

We know that this limit exists, and thus

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

by the definition of lipschitz and the limit.

Conversely, if

$$|f'(x)| \leq M \quad \forall x \in \mathbb{R}$$

we want

$$|f(x) - f(y)| \leq M|x - y| \quad \forall x, y \in \mathbb{R}$$

By the MVT, $\exists c \in (x, y)$ such that

$$f(x) - f(y) = f'(c)(x - y)$$

Then

$$|f(x) - f(y)| = |f'(c)||x - y| \leq M|x - y|$$
$$|f(x) - f(y)| \leq M|x - y|$$

$\square$

Definition: Let $T : X \to X$ be a <u>contraction</u> if $\exists\, 0 \le c < 1$ such that $d(T(p), T(q)) \le c \cdot d(p, q)\ \ \forall p, q \in X$. Basically, imagine a lipschitz function with $M < 1$, and all the contraction does is "minimize" the function by a factor of c.

Theorem: big boy; Let X be a complete (reminder, this means that all cauchy sequences converge) metric space, and $T : X \to X$ a contraction. Then this means $\exists$ unique $p \in X$ such that $T(p) = p$, called a <u>fixed point</u>

Proof:

- Uniqueness: Suppose we have two fixed points, eg $T(p_1) = p_1, T(p_2) = p_2$. Then

$$d(p_1, p_2) = d(T(p_1), T(p_2)) \le c \cdot d(p_1, p_2)$$

which is a contradiction unless $d(p_1, p_2) = 0$. Thus, $p_1 = p_2$.

- Existence: Pick any $p_0 \in X, p_{k+1} = T(p_k)$. We know that for some $p_0, p_1, p_2, \ldots$. We will show that $\{p_k\}$ is cauchy. Then

$$d(p_1, p_0) \le c \cdot d(p_1, p_0)$$
$$d(p_1, p_2) = d(T(p_0), T(p_1)) \le c \cdot d(p_0, p_1)$$
$$d(p_2, p_3) = d(T(p_1), T(p_2)) \le c \cdot d(p_1, p_2) \le c^2 \cdot d(p_0, p_1)$$
$$\vdots$$
$$d(p_k, p_{k+1}) \le c^k \cdot d(p_0, p_1)$$

Thus

$$\begin{aligned}
d(p_k, p_{k+l}) &\le d(p_{k+l}, p_{k+l-1}) + d(p_{k+l-1}, p_k) \\
&\le d(p_{k+l}, p_{k+l-1}) + d(p_{k+l-1}, p_{k+l-2}) + \cdots + d(p_{k+1}, p_k) \\
&\le d(p_{k+l}, p_{k+l-1}) + \cdots + d(p_{k+1}, p_k) \\
&\le (c^{k+l-1} + c^{k+l-2} + \cdots + c^k) \cdot d(p_0, p_1) \\
&\le (c^k \sum_{l=0}^{\infty} c^l) d(p_0, p_1) \\
&= c^k (\frac{1}{1-c}) d(p_0, p_1) \\
&\to 0 \text{ uniformly as } k \to \infty
\end{aligned}$$

Now, let $\{p_k\} \to p$. We will show that $T(p) = p$. We must show that T is continuous, i.e. if $p_k \to p$, then $T(p_k) \to T(p)$. Then

$$d(T(p_k), T(p)) \le C \cdot d(p_k, p) \to 0 \text{ basically, the lipschitz condition}$$

And thus we have that $\{T(p_k)\} \to T(p)$.

We now know that $\{p_k\} \to p, \{T(p_k)\} \to T(p)$. Then $\{p_{k+1}\} \to T(p)$, and $\{p_{k+1}\} \to p$ by the uniqueness of limits. Thus, $T(p) = p$ $\square$

Set Up: Suppose we have a set $O \subseteq \mathbb{R}^2$, open. Then let us have a point $(x_0, y_0) \in O$. Let there be a function $g : O \to \mathbb{R}$, continuous. Let I be an interval, $f : I \to \mathbb{R}$ such that the graph of $\{(x, f(x)) | x \in I\}$ is contained in O.
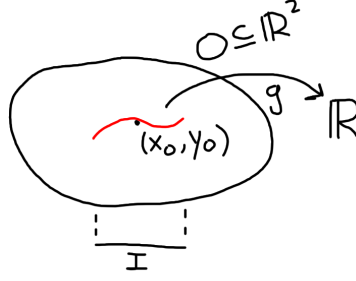
Figure 1: The setup (krita lol)

The Existence Theorem: Let $\mathcal{O}$ be an open subset of the plane $\mathbb{R}^2$ that contains the point (x_0, y_0). Suppose that the function $g : \mathcal{O} \to \mathbb{R}^2$ is continuous and that there's a positive number M such that

$$|g(x, y_1) - g(x, y_2) \le M|y_1 - y_2| \quad \forall (x, y_1), (x, y_2) \in \mathcal{O}$$

Then, there exists a unique function $f : I \to \mathbb{R}$ such that

$$\begin{cases} f'(x) = g(x, f(x)) \quad \forall x \in I \\ f(x_0) = y_0 \end{cases}$$

Proof: For l, a positive number, define I_l to be the closed interval $[x_0 - l, x_0 + l]$. We will show that l can be chosen such that there is exactly one continuous function $f : I_l \to \mathbb{R}$ that satisfies

$$f(x) = y_0 + \int_{x_0}^{x} g(s, f(s))ds \quad \forall x \in I_l$$

Once such an l is chosen, it follows from the Equivalence Lemma that there is exactly one solution of the differential equation on the interval $I = (x_0 - l, x_0 + l)$ that satisfies the initial condition $f(x_0) = y_0$.

Since $\mathcal{O}$ is open, we can choose positive numbers a, b such that the rectangle $R = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$ is contained in $\mathcal{O}$. Now for each positive number l with $l \le a$, define X_l to be the subspace of the metric space $C(I_l, \mathbb{R})$ consisting of the continuous functions $f : I_l \to \mathbb{R}$ with the property that

$$|f(x) - y_0| \le b \quad \forall x \in I_l$$

Observe that X_l consists of the continuous functions $f : I_l \to \mathbb{R}$ whose graphs are contained in the rectangle R.

For a function $f \in X_l$, define $T(f) \in C(I_l, \mathbb{R})$ by

$$T(f)(x) = y_0 + \int_{x_0}^{x} g(t, f(t))dt \quad \forall x \in I_l$$

Observe that a solution of the integral equation is simply a fixed point of the mapping $T : X_l \to C(I_l, \mathbb{R})$.

Since $C(I_l, \mathbb{R})$ is a complete metric space and X_l is a closed subset of it, it follows from a Corollary that X_l is also a complete metric space. Now we will show that if l is chosen to be sufficiently small, then

$$T(X_l) \subseteq X_l, T : X_l \to X_l \text{ is a contraction}$$

and from the Contraction Mapping Principle it will follow that there is a unique fixed point of T in X_l, which is a solution of the integral equation. Choose some positive K such that

$$|g(x, y)| \leq K \quad \forall (x, y) \in R$$

, which is allowed by the EVT. If f is a function in X_l and the point x belongs to I_l, then

$$|T(f)(x) - y_0| = |\int_{x_0}^{x} g(t, f(t)) dt|$$
$$\leq l \cdot K$$

and thus $T(X_l) \subseteq X_l$ assuming that $l \cdot K \leq b$. Now observe that if $f_1, f_2 \in X_l, x \in I_l$, then

$$|T(f_1)(x) - T(f_2)(x)| = |\int_{x_0}^{x} g(t, f_1(t)) - g(t, f_2(t)) dt|$$
$$\leq |\int_{x_0}^{x} M \cdot d(f_1(t), f_2(t)) dt$$
$$\leq |x - x_0| \cdot M \cdot d(f_1, f_2)$$
$$\leq l \cdot M \cdot d(f_1, f_2)$$

and thus $T : X_l \to X_l$ is a contraction. It follows that there is a unique fixed point of T in X_l, which is a solution of the integral equation. $\square$

Lemma: Assuming this set up, TFAE:

- f is differentiable and

$$\begin{cases} f'(x) = g(x, f(x)) & \forall x \in I \\ f(x_0) = y_0 \end{cases}$$

- f is continuous and $f(x) = y_0 + \int_{x_0}^{x} g(t, f(t)) dt$

Proof: Assume 1. Then f diff $\implies f$ continuous $\implies f'$ continuous, since $f'(x) = g(x, f(x))$. Then

$$f(x) - y_0 = f(x) - f(x_0) = \int_{x_0}^{x} f'(t) dt = \int_{x_0}^{x} g(t, f(t)) dt$$

which is 2.

Assume 2. Since $f(x) = y_0 + \int_{x_0}^{x} g(t, f(t)) dt$, f is differentiable and $f'(x) = g(x, f(x))$. Also, $f(x_0) = y_0$ can be seen by taking $x = x_0$. Thus 1. $\square$

Theorem: Assume, in addition, that $\exists M$ such that $|g(x, y_1) - g(x, y_2)| \leq M|y_1 - y_2| \quad \forall (x, y_1), (x, y_2) \in O$. Let $a, b > 0$ s.t.

$$[x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b] \subseteq O$$

(imagine a box around the point in the illustration above). Then define, for $0 < l \leq a$,

$$X_l = \{f : [x_0 - l, x_0 + l] \to [y_0 - b, y_0 + b]\}$$

Then $\exists l > 0$ such that $T(f)(x) = y_0 + \int_{x_0}^{x} g(t, f(t)) dt$ is a mapping $T : X_l \to X_l$, and is a contraction. Then $\exists! f \in X_l$ such that

$$f(x) = y_0 + \int_{x_0}^{x} g(t, f(t)) dt \text{ ,and}$$
$$f'(x) = g(x, f(x)) \text{ ,and}$$
$$f(x_0) = y_0$$

Proof:

$$\exists K \text{ s.t. } |g(x,y)| \leq K \quad \forall (x,y) \in [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$$

We need to show that $T : X_l \to X_l$ if l is sufficiently small. Consider $|T(f)(x) - y_0|$. Then

$$
\begin{aligned}
|T(f)(x) - y_0| = |\int_{x_0}^x g(t, f(t)) dt| \\
\leq |x - x_0| \cdot \max_{t \in [x_0 - l, x_0 + l]} |g(t, f(t))| \\
\leq |x - x_0| \cdot K \\
\leq lK
\end{aligned}
$$

Thus if l is really small, then $T(f)(x) \in [y_0 - b, y_0 + b]$. Thus $T : X_l \to X_l$.

Now we need to demonstrate that this is a contraction, provided that l is sufficiently small. Then $\exists l > 0$ such that

$$T(f)(x) = y_0 + \int_{x_0}^x g(t, f(t)) dt$$

We want that

$$d(T(f_1), T(f_2)) \leq c \cdot d(f_1, f_2)$$

Then

$$
\begin{aligned}
|T(f_1)(x) - T(f_2)(x)| \leq \int_{x_0}^x |g(t, f_1(t)) - g(t, f_2(t))| dt \\
\leq \int_{x_0}^x M |f_1(t) - f_2(t)| dt \\
\leq M \cdot l \cdot \max_{t \in [x_0 - l, x_0 + l]} |f_1(t) - f_2(t)| \\
= M \cdot l \cdot d(f_1, f_2)
\end{aligned}
$$

Then we get $d(T(f_1), T(f_2)) \leq M \cdot l \cdot d(f_1, f_2)$. So as long as $M \cdot l < 1$, then T is a contraction.

Also, there thus $\exists! f : (x_0 - l, x_0 + l) \to [y_0 - b, y_0 + b]$ such that

$$
\begin{aligned}
f(x) = y_0 + \int_{x_0}^x g(t, f(t)) dt \text{ ,and} \\
f'(x) = g(x, f(x)) \text{ ,and} \\
f(x_0) = y_0
\end{aligned}
$$

$\square$

Note: come back and review this;

Recall: For a setup $B = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$, $g : B \to \mathbb{R}$ continuous, such that $\exists M$ such that

$$|g(x, y_1) - g(x, y_2)| \leq M |y_1 - y_2| \quad \forall (x, y_1), (x, y_2) \in B$$

Theorem: If g is as above, $\exists 0 < l < a$ and

$$f : [x_0 - l, x_0 + l] \to [y_0 - b, y_0 + b] \text{ where b is such that } B = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b] \subseteq O$$

, continuous, such that

$$f(x) = y_0 + \int_{x_0}^{x} g(t, f(t))dt \quad \forall |x - x_0| \leq l$$

and thus

$$f'(x) = g(x, f(x)) \quad \forall |x - x_0| \leq l, f(x_0) = y_0$$

In addition, if $f_1, f_2 : [x_0 - l, x_0 + l] \to [y_0 - b, y_0 + b]$ are continuous, and satisfy the above conditions, then $f_1 = f_2$.

If we drop the lipschitz condition, then we can't guarantee uniqueness.

<u>Remark:</u> We can find an example of continuous g for which uniqueness fails given no lipschitz condition. Take for instance take a function that's very close to 0, and then goes up after 0.

$$f(t) = \begin{cases} 0 & \text{if } t < 0 \\ t^3 & \text{if } t \geq 0 \end{cases}$$

Then,

$$f'(x) = \begin{cases} 0 & \text{if } x < 0 \\ 3x^2 & \text{if } x \geq 0 \end{cases}$$

In other words $f(t)$ is a solution for the first order ODE $f'(t) = 3f(t)^{2/3}$, and $f'(0) = 0$. This is one solution. We claim this solution isn't unique. Another solution is $f(t) = 0$ for all t.

In fact for the format

$$f(t) = \begin{cases} 0 & \text{if } t < 0 \\ (t - c)^3 & \text{if } t \geq 0 \end{cases}$$

has infinitely many solutions. To demonstrate that this is not lipschitz, let's give that $g(t, y) = 3y^{2/3}$, and g is continuous. Then

$$|g(t, y_1) - g(t, y_2)|?M|y_1 - y_2|$$

The derivative of the function is not bounded near 0, and thus the lipschitz condition fails, eg $\frac{\partial g}{\partial y}(t, y)$ is unbounded as $y \to 0$.

<u>Remark:</u> Existence is a local property. As an example,

$$\begin{cases} f'(t) = f^2(t) \\ f(0) = 1 \end{cases}$$

Then $f(t) = \frac{1}{1-t}$ is a solution for $(-\infty < t < 1)$. But existence is not a global property. For instance, $f(t) = \frac{1}{1-t}$ is not necessarily a solution for $t \geq 1$.

Theorem: Let $g : \mathbb{R}^2 \to \mathbb{R}$, and assume that

$$|g(t, y_1) - g(t, y_2)| \leq M|y_1 - y_2| \quad \forall (t, y_1), (t, y_2) \in \mathbb{R}^2, M > 0$$

If

$$f_1'(t) = g(t, f_1(t)), f_1(t_0) = y_0$$
$$f_2'(t) = g(t, f_2(t)), f_2(t_0) = y_0 \quad \forall t \in \mathbb{R}$$

Then $f_1(t) = f_2(t)$ for all $t \in \mathbb{R}$. We will show this two ways, one in a sort of diff-eq way and one in a topology way.

Proof 1: (The diff-eq proof) We will show that $f_1(t) = f_2(t) \quad \forall t \geq t_0$. A similar proof will work for $t < t_0$. Consider the point

$$E(t) = (f_1(t) - f_2(t))^2$$

Machedon remarks he uses E because he thinks of it as energy

Then

$$\begin{aligned}
E'(t) &= 2(f_1(t) - f_2(t))(f_1'(t) - f_2'(t)) \\
&\leq 2|f_1(t) - f_2(t)||g(t, f_1(t)) - g(t, f_2(t))| \\
&\leq 2M|f_1(t) - f_2(t)||f_1(t) - f_2(t)| \\
&= 2ME(t)
\end{aligned}$$

Thus, $E'(t) \leq 2ME(t), E(t_0) = 0$. Then $E'(t) - 2ME(t) \leq 0$. Using an integrating factor, we will solve this;

$$\frac{d}{dt}(e^{-2Mt}E(t)) = e^{-2Mt}E'(t) - 2Me^{-2Mt}E(t) \leq 0$$

We know that $E(t)e^{-2Mt}$ is decreasing, it is 0 at t_0. We would like to conclude it's 0. We know that $E(t)$ is positive, and also e^{-2Mt} is positive. Thus $E(t) = 0$ for all $t \geq t_0$. Thus $f_1(t) = f_2(t)$ for all $t \geq t_0$. A similar proof will work for $t < t_0$. □

Proof 2: (The topology proof) Let $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ is differentiable. Then

$$\begin{aligned}
f_1'(t) &= g(t, f_1(t)), f_1(t_0) = y_0 \\
f_2'(t) &= g(t, f_2(t)), f_2(t_0) = y_0 \quad \forall t \in \mathbb{R}
\end{aligned}$$

, etc, and also Lipschitz

$$|g(t, y_1) - g(t, y_2)| \leq M|y_1 - y_2| \quad \forall (t, y_1), (t, y_2) \in \mathbb{R}^2, M > 0$$

From our main theorem last time, we know that

$$\exists l > 0 \text{ s.t. } f_1(t) = f_2(t) \quad \forall |t - t_0| \leq l$$

Consider the set

$$S = \{t | f_1(t) = f_2(t)\}$$

given that f_1, f_2 continuous. Then the set S is closed. Let $t_1 \in S$. Then $f_1(t_1) = f_2(t_1)$. We claim that S is also open. Then, $\exists \delta > 0$ such that $f_1(t) = f_2(t) \quad \forall |t - t_1| < \delta$ because they satisfy the same diff'eq, and they take the same value. We claim that there's a neighborhood of this point where they are still the same. There's a theorem (local uniqueness theorem) applied to t_1 that says that there's a neighborhood of t_1 where the solution is unique. Thus, S is open. Since S is both open and closed, $S = \mathbb{R}$ and thus $f_1 = f_2$. (It's not empty because t_0) □

7.2 Unnamed Topic

<u>Remark:</u> In the metric space $C([0, 1], \mathbb{R})$, with $\|f\| = \max_{x \in [0,1]} |f(x)|$, closed and bounded sets need not be sequentially compact. As an example, take the closed unit ball $B = \{f \in C([0, 1], \mathbb{R}) | \|f\| \leq 1\}$, and $f_k(x) = x^k$ is a sequence in B. We claim that this sequence has no convergent subsequence.

This is because the limit of the sequence is not continuous.

$$f_k(x) \to \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$

The limit of such a convergence subsequence would have to be continuous, and thus this sequence has no convergent subsequence.

7.3   Reviewing Exam Problems

Problem 39a: Let

$$X = \{f \in C([0,\pi], \mathbb{R}), 1 - \pi \le f(x) \le 1 + \pi \ \ \forall x \in [0,\pi]\}$$

$$T : X \to C([0,\pi], \mathbb{R}), T(f)(x) = 1 + \int_0^x \cos(f(t))dt$$

Prove $T : X \to X$.

Answer: Let $f \in X$. We have to prove $1 - \pi \le T(f(x)) \le 1 + \pi \ \ \forall x \in [0,\pi]$. Then

$$1 - \pi \le T(f(x)) \le 1 + \pi$$

$$1 - \pi \le 1 + \int_0^x \cos(f(t))dt \le 1 + \pi$$

$$-\pi \le \int_0^x \cos(f(t))dt \le \pi$$

i.e. we have to show that the integral is bounded by π. Then

$$\left| \int_0^x \cos(f(t))dt \right| \le \int_0^x |\cos(f(t))|dt \le \int_0^x 1 dt = x \le \pi$$

Problem 39b: Is T a contraction?

Answer: Let $f_1, f_2 \in X$. Then

$$(T(f_1)(x)) = 1 + \int_0^x \cos(f_1(t))dt$$

$$(T(f_2)(x)) = 1 + \int_0^x \cos(f_2(t))dt$$

and the distance

$$d(T(f_1), T(f_2)) = \max_{x \in [0,\pi]} |T(f_1)(x) - T(f_2)(x)|$$

$$= \max_{x \in [0,\pi]} \left| \int_0^x \cos(f_1(t)) - \cos(f_2(t))dt \right|$$

$$\le \int_0^\pi |\cos(f_1(t)) - \cos(f_2(t))|dt$$

$$\le C \cdot d(f_1, f_2)$$

The inside of the last integral is equal by the MVT to $\sin(\theta(t))|f_1(t) - f_2(t)|$, and thus the integral is bounded by $\pi \cdot d(f_1, f_2)$. Thus, T is a contraction. To elaborate the MVT;

$$g(x) - g(y) = g'(\theta)(x - y)$$

25

for some $\theta \in (x, y)$.

Try

$$f_1(t) = 0$$
$$f_2(t) = \pi$$

Then $d(f_1, f_2) = \pi$, and

$$T(f_1) = 1 + \int_0^x \cos(0) dt = 1 + x$$
$$T(f_2) = 1 + \int_0^x \cos(\pi) dt = 1 - x$$
$$d(T(f_1), T(f_2)) = \max_{x \in [0, \pi]} |1 + x - (1 - x)| = 2\pi$$

So the Lipschitz constant is at least 2. For this specific f_1, f_2, $d(T(f_1), T(f_2)) = 2\pi \cdot d(f_1, f_2)$. So it is not a contraction (not < 1).

Problem 39c: If f is a fixed point of T, what differential equation does f satisfy?

Answer:

$$f(x) = 1 + \int_0^x \cos(f(t)) dt$$
$$f'(x) = \cos(f(x)), f(0) = 1$$

Does the solution to this exist for all $x \geq 0$? Then

$$g(x, y) = \cos(y), |g(x, y)| \leq K = 1$$
$$|g(x, y_1) - g(x, y_2)| \leq M|y_1 - y_2|, M = 1$$

Thus there exists $h = h(K, M)$ such that the solution to $f'(x) = \cos(f(x)), f(0) = 1$ exists and is unique for $[x_0 - h, x_0 + h]$. Thus the solution exists for all $x \geq 0$ by just splitting it up into intervals of length h.

Remark: Let $g : \mathbb{R}^2 \to \mathbb{R}$, continuous, such that

$$\exists M, K \text{ with } |g(x, y)| \leq K \quad \forall (x, y) \in \mathbb{R}^2$$
$$|g(x, y_1) - g(x, y_2)| \leq M|y_1 - y_2| \quad \forall (x, y_1), (x, y_2) \in \mathbb{R}^2$$

Then the solution to $f'(x) = g(x, f(x)), f(x_0) = y_0$ exists and is unique for all $x \in \mathbb{R}$. This works because the solution exists for $[x_0 - h, x_0 + h]$ for some h, and then we can just keep extending the interval - and x_0 is arbitrary (eg K, M are independent of x_0, y_0).

Problem 40a: Let

$$X = \{f \in C([0, 1], \mathbb{R}), 0 \leq f(x) \leq 1 \quad \forall x \in [0, 1]\}$$
$$(T(f))(x) = \int_0^x f(t)^{\frac{2}{3}} dt$$

Prove that $T : X \to X$.

Answer: If $0 \leq f \leq 1$, then $0 \leq f^{\frac{2}{3}} \leq 1$. Then

$$0 \leq \int_0^x f(t)^{\frac{2}{3}} dt \leq \int_0^x 1 dt = x \leq 1$$

Thus $T : X \to X$.

Problem 40b: Is T a contraction?

Answer: We expect T not to be a contraction because the integral is not bounded by a constant. Let's try $f_1(t) = 0, f_2(t) = 1$. Then

$$T(f_1)(x) = 0$$

$$T(f_2)(x) = \int_0^x 1 dt = x$$

Then $d(T(f_1), T(f_2)) = 1 \cdot d(f_1, f_2)$. Thus T is not a contraction.

Is it the case that using a sufficiently small f_2 will make it a contraction? eg. does there exist $h > 0$ such that $T : X \to X$ is a contraction?

Let

$$f_1(t) = 0$$
$$f_2(t) = \delta(\text{small})$$

Then $d(f_1, f_2) = \delta$. Then

$$T(f_1)(x) = 0$$

$$T(f_2)(x) = \int_0^x \delta^{\frac{2}{3}} dt = \delta^{\frac{2}{3}} x$$

Then $d(T(f_1), T(f_2)) = h \cdot \delta^{\frac{2}{3}}$. Does there exist a C such that that this is less than or equal to $C \cdot \delta$? No.

Problem something: Let $X = \{f \in C([0,1], \mathbb{R}), |f(x)| \leq 1\}$, and $T : X \to C([0,1], \mathbb{R})$ be

$$T(f)(x) = \int_0^x \cos(f(t)) dt$$

Show that $T : X \to X$.

Answer: We check that

$$|\int_0^x \cos(f(t)) dt| \leq 1 \quad \forall x \in [0,1]$$

$$\leq \int_0^1 |cos(f(t))| dt \leq 1$$

so $T : X \to X$.

Problem something else: Is T a contraction?

Answer:

$$|T(f_1)(x) - T(f_2)(x)| = |\int_0^x \cos(f_1(t)) - \cos(f_2(t)) dt|$$

$$?C \cdot \max_{t \in [0,1]} |f_1(t) - f_2(t)|$$

Is this true? Although moving abs val inside a definite integral is not always true, we can get away with it

27

for this one

$$\left|\int_0^x \cos(f_1(t)) - \cos(f_2(t))dt\right| \le \int_0^1 |\cos(f_1(t)) - \cos(f_2(t))|dt$$

$$= \int_0^1 |\sin(\theta(t))||f_1(t) - f_2(t)|dt \text{ for some } \theta(t)$$

$$\le \int_0^1 |f_1(t) - f_2(t)|dt$$

However we missed finding a c such that this is less than or equal to $c \cdot \max_{t \in [0,1]} |f_1(t) - f_2(t)|$. Thus this method doesn't work (since we found $c = 1$). But it kind of does; let's continue

$$\le \sin(1) \cdot \max_{t \in [0,1]} |f_1(t) - f_2(t)|$$

and since $\sin(1) < 1$, we have that T is a contraction.

Problem something else else: Let $X = \{f \in C([0, \frac{1}{2}],, 0 \le f(x) \le 4)\}$, and $T : X \to C([0, \frac{1}{2}], \mathbb{R})$ be

$$T(f)(x) = 1 + \int_0^x f(t)dt$$

$T : X \to X$?

Answer:

$$0 \le 1 \le T(f)(x) \le 1 + \int_0^x f(t)dt$$

$$\le 1 + \int_0^{\frac{1}{2}} 4dt \le 3 \le 4$$

and thus $T : X \to X$.

Problem something else else else: Is T a contraction?

$$|T(f_1)(x) - T(f_2)(x)| = \left|\int_0^x f_1(t) - f_2(t)dt\right|$$

$$\le \int_0^{\frac{1}{2}} |f_1(t) - f_2(t)|dt$$

$$\le \frac{1}{2} \cdot \max_{t \in [0, \frac{1}{2}]} |f_1(t) - f_2(t)|$$

so T is a contraction.

Problem something else else else else: Does it have a fixed point?

Answer:

$$f(x) = 1 + \int_0^x f(t)dt \implies f'(x) = f(x), f(0) = 1$$

Thus $f(x) = e^x$ is a solution and is unique. Thus T has a fixed point.

Problem 1: Let $f : \mathbb{R} \to \mathbb{R}$. Then $\exists \delta > 0$ such that $|f(x) - f(y)| \le \epsilon \;\; \forall |x - y| \le \delta$. What does this mean?

Answer: Well, it's uniformly continuous, but that's not the point; it means that f is constant, since it implies $|f(x) - f(y)| = 0$ for all x, y.

Problem 2a:

$$\max_{(x,y) \neq (0,0)} \frac{x + 2y}{\sqrt{x^2 + y^2}} = \max_{(x,y) \neq (0,0)} \frac{\langle (x,y), (1,2) \rangle}{\|(x,y)\|} \le \|(1,2)\| = \sqrt{5}$$

Then the maximizer of (x, y) is $(1, 2)$

Problem 2b:

$$\max_{(x,y) \neq (0,0)} \frac{x + y}{\sqrt{x^2 + 4y^2}} = \max_{(x,y) \neq (0,0)} \frac{\langle (x, 2y), (1, \frac{1}{2}) \rangle}{\|(x, 2y)\|} \le \|(1, \tfrac{1}{2})\| = \sqrt{\frac{5}{4}}$$

Then the maximizer of (x, y) is $(1, \frac{1}{4})$ since we want $(x, 2y) = (1, \frac{1}{2})$.

Problem 3: $f : [1, \infty) \to \mathbb{R}, f(x) = \frac{1}{x}$. Is this uniformly continuous?

Answer: We need to find

$$|f(x_1) - f(x_2)| \le 1 \cdot |x_1 - x_2| \;\; \forall x_1, x_2 \in [1, \infty)?$$

$$|f(x_1) - f(x_2) \le |f'(\theta)||x_1 - x_2| = \frac{1}{\theta^2}|x_1 - x_2| \le 1 \cdot |x_1 - x_2|$$

so f is uniformly continuous.

Problem 4: Lemma: If $f : S \subseteq \mathbb{R} \to \mathbb{R}$ is uniformly continuous if it is lipschitz continuous on S.

Answer: We know that $\exists M$ s.t. $|f(x_1) - f(x_2)| \le M \cdot |x_1 - x_2| \;\; \forall x_1, x_2 \in S$. Then we want to find

$$\forall \epsilon > 0 \exists \delta > 0 \text{ s.t. } |x_1 - x_2| \le \delta \implies |f(x_1) - f(x_2)| \le \epsilon$$

Fix $\epsilon > 0, \delta = \frac{\epsilon}{M}$ works. Then

$$|x_1 - x_2| \le \delta \implies |f(x_1) - f(x_2)| \le M \cdot |x_1 - x_2| \le M \cdot \delta = \epsilon$$

and thus f is uniformly continuous.

Problem 5: Let $f_k = x^k(1 - x)^k, f_k \in C([0, 1], \mathbb{R})$. Is f_k Cauchy? Yes, it converges uniformly to 0.

8 Differentiating Multivar Functions

Chapter 13

8.1 Limits

Definition: x_* is a limit point of $A \subseteq \mathbb{R}^n$ if

$$\exists \{x_k\} \subseteq A \setminus \{x_*\} \text{ s.t. } \{x_k\} \to x_*$$

Definition: If $f : A \to \mathbb{R}$, and x_* is a limit point of A, then

$$\lim_{x \to x_*} f(x) = L \text{ if } \forall \{x_k\} \subseteq A \setminus \{x_*\},$$
$$\lim_{k \to \infty} f(x_k) = L$$

Suppose

$$f(x,y) = \frac{xy}{x^2 + y^2}, f : \mathbb{R}^2 \setminus \{(0,0)\}$$

. Then the limit

$$\lim_{(x,y)\to(0,0)} f(x,y) \text{ DNE}$$

To show this, consider $x_k = (\frac{1}{k}, 0) \to (0,0)$. Then $f(\frac{1}{k}, 0) = 0 \ \forall k$. So if the limit does exist, then it has to be 0. Take $x_k = (\frac{1}{k}, \frac{1}{k})$. Then $f(\frac{1}{k}, \frac{1}{k}) = \frac{1}{2} \ \forall k$. Thus the limit does not exist.

<u>Example:</u> Let

$$g(x,y) = \frac{x^2 y}{x^2 + y^2}$$

We claim that

$$\lim_{(x,y)\to(0,0)} g(x,y) = 0$$

Let (x_k, y_k) be a sequence of points in $\mathbb{R}^2$ that approach $(0,0)$. Then

$$|g(x_k, y_k)| \leq |x_k| \frac{|x_k y_k|}{x_k^2 + y_k^2}$$

The right half of the right side is bounded by $\frac{1}{2}$, and the left half goes to 0. Thus the limit is 0.

Definition: $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is homogeneous of degree k if

$$f(tx) = t^k f(x) \ \forall t > 0, x \in \mathbb{R}^n \setminus \{0\}$$

<u>Remark:</u> What's special about f, g? According to the definition, g is homogeneous of degree 1. f is homogeneous of degree 0. Note that g had a limit, and f did not. Is this generalizable in some way?

Proposition: If $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is continous and homogeneous of degree $k > 0$, then

$$\lim_{x \to 0} f(x) = 0$$

Proof: Let

$$f(x) = f(\|x\| \cdot \frac{x}{\|x\|})$$
$$= \|x\|^k f(\frac{x}{\|x\|})$$

We try to bound the right side; consider the unit sphere, which is sequentially compact. We thus notice that $\frac{x}{\|x\|} \in S^{n-1}$, which is the $n-1$ dimensional sphere, and is seq. compact. We know that f is continuous, so $f(\frac{x}{\|x\|})$ is bounded. Then for some M

$$|f(x)| = |\|x\|^k f(\frac{x}{\|x\|})|$$
$$\leq \|x\|^k \cdot M$$

Then $\lim_{x\to 0} f(x) = 0$.

<u>Remark:</u> What about functions homogeneous with degree 0? We can't say anything about them. For instance, all constant functions are homogeneous of degree 0, and they can have a limit at 0.

8.2 Partial Derivatives

Definition: Let $f : O \to \mathbb{R}$, $x = (x_1, x_2, ..., x_n) \in O \subseteq \mathbb{R}^n$, O open. Then

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t}$$

, if the limit exists. i.e. we keep all of the variables other than x_i fixed, and then take $\frac{d}{dx_i}$.

Example: Let $f : \mathbb{R}^2 \to \mathbb{R}$, and

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

We notice that f is not continuous at $(0, 0)$. Note that $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ exist by the quotient rule at all $(x, y) \neq (0, 0)$. As for the partial of f wrt x at 0,

$$\lim_{t \to 0} \frac{f(t, 0) - f(0, 0)}{t} = 0$$

Definition: If one $\frac{\partial f}{\partial x_i}(x)$ exists $\quad \forall x \in O$, define

$$\frac{\partial f}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial^2 x}{\partial x_i x_j}$$

Theorem: This is the theorem that says that if the partials exist and are continuous, then the order of differentiation doesn't matter. But I was mentally tapped out and didn't write it; I'll come back to it later. In the book, this is Thm 13.10

Remark: In general

$$\frac{d}{dt}\Big|_{t=0}[f(x + th)] = \langle \nabla f(x), h \rangle$$

where $\nabla f(x)$ is the gradient of f at x.

Proposition: Let $f : \mathbb{R}^n \to \mathbb{R}$, and assume that all $\frac{\partial f}{\partial x_i}(x)$ exists $\quad \forall x \in \mathbb{R}^n$, $\forall i \in \{1, \ldots, n\}$. Fix $x, h \neq 0 (h \in \mathbb{R}^n)$. Then $\exists z_1, \ldots, z_n \in B_{\|h\|}(x)$ such that

$$f(x + h) - f(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(z_i)h_i$$

Proof: Textbook has $n = 3$. For us, let $n = 2$. Then

$$\begin{aligned} f(x + h) - f(x) &= f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2) \\ &= f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) + f(x_1, x_2 + h_2) - f(x_1, x_2) \\ &= \frac{\partial f}{\partial x_1}(x_1 + \theta_1 h_1, x_2 + h_2)h_1 + \frac{\partial f}{\partial x_2}(x_1, x_2 + \theta_2 h_2)h_2 \text{ for some } \theta_1, \theta_2 \in (0, 1) \end{aligned}$$

by MVT. Then

$$f(x + h) - f(x) = \sum_{i=1}^{2} \frac{\partial f}{\partial x_i}(z_i)h_i$$

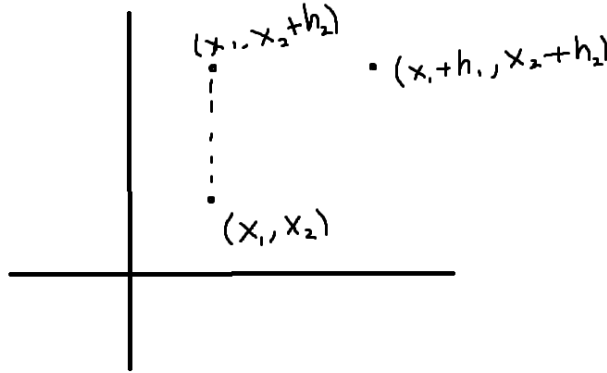for some $z_1, z_2 \in B_{\|h\|}(x)$.

Figure 2: illustration

Definition: If $f : O \to \mathbb{R}$, $O \subseteq \mathbb{R}^n$ open, and all $\frac{\partial f}{\partial x_i}$ exist and are continuous, then f is called continuously differentiable, or C^1.

Theorem: If $f : \mathbb{R}^n \to \mathbb{R}$ is C^1, then $\quad \forall x \in \mathbb{R}^n, h \neq 0$, the limit

$$\frac{d}{dt}|_{t=0} f(x + th) = \lim_{t \to 0} \frac{f(x + th) - f(x)}{t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i = \langle \nabla f(x), h \rangle$$

Proof: Consider $\frac{f(x+th) - f(x)}{t}$. By our previous proposition, this is

$$\frac{\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(z_i) t \cdot h_i}{t}$$

for some $z_1, \ldots, z_n \in B_{t\|h\|}(x)$. Then as t approaches 0 all of the z_i approach x, and thus the limit is

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i$$

$\square$

8.3 Directional Derivative

Definition: Let $p \neq 0, p \in \mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R}$. The directional derivative of f at x in the direction of p is

$$\frac{\partial f}{\partial p}(x) = \frac{d}{dt}|_{t=0} f(x + tp) \text{ if it exists}$$

and recall that

$$\nabla f(x) = (\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x))$$

and thus what we proved in the theorem immediately preceding this subsection is that if $f \in C^1$, then

$$\frac{\partial f}{\partial p}(x) = \langle \nabla f(x), p \rangle$$

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}$, C^1, $x \in \mathbb{R}^n, h \in \mathbb{R}^n, h \neq 0$. Then

$$\exists 0 < \theta < 1 \text{ s.t. } f(x + h) - f(x) = \langle \nabla f(x + \theta h), h \rangle$$

Proof: Let $\phi : \mathbb{R} \to \mathbb{R}, \phi(t) = f(x + th)$ so that $\phi(1) = f(x + h), \phi(0) = f(x)$. Then

$$f(x + h) - f(x) = \phi(1) - \phi(0) = \phi'(\theta) = \langle \nabla f(x + \theta h), h \rangle$$

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}$, and assume that all $\frac{\partial f}{\partial x_i}(x)$ exist and are continuous. Then f is continuous.

Proof: Consider $|f(x + h) - f(x)|$. By the previous theorem, this is

$$|\langle \nabla f(x + \theta h), h \rangle| \leq \|\nabla f(x + \theta h)\| \|h\|, \theta \in (0, 1)$$
$$\to 0 \text{ as } h \to 0$$

and thus f is continuous.

Remark: What's a slightly weaker condition than the partial derivatives being continuous that would still ensure continuity of f? We can say that f is C^1 if all of the partial derivatives exist and are bounded.

You can see this in the latter half of the inequality; we just need $\|\nabla f(x + \theta h)\|$ to be bounded.

And if

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

then

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^2 + y^2) - xy(2x)}{(x^2 + y^2)^2}$$
$$= \frac{y^3 - x^2 y}{(x^2 + y^2)^2}$$

and

$$\frac{\partial f}{\partial y}(x, y) = \frac{x(x^2 + y^2) - xy(2y)}{(x^2 + y^2)^2}$$
$$= \frac{x^3 - xy^2}{(x^2 + y^2)^2}$$

they are both homogeneous of degree -1, but are unbounded.

Remark: Let $f : \mathbb{R}^n \to \mathbb{R}, C^1$. Fix x Then

$$\max_{\|p\|=1} \frac{\partial f}{\partial p}(x)$$

is attained for $p = \frac{\nabla f(x)}{\|\nabla f(x)\|}$.

Proof: If $\|p\| = 1$, then

$$\frac{\partial f}{\partial p}(x) = \langle \nabla f(x), p \rangle$$
$$\leq \|\nabla f(x)\| \|p\|$$
$$\leq \|\nabla f(x)\|$$

and thus the maximum would be $\|\nabla f(x)\|$. Now we show that this is attained if $p = \frac{\nabla f(x)}{\|\nabla f(x)\|}$. Then

$$\frac{\partial f}{\partial p}(x) = \langle \nabla f(x), \frac{\nabla f(x)}{\|\nabla f(x)\|} \rangle$$
$$= \|\nabla f(x)\|$$

$\square$

Remark: Let $f : \mathbb{R}^n \to \mathbb{R}, C^1$. Then

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

Proof: The left hand side;

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

We do this in order to define differentiable functions.

Definition: For $f : \mathbb{R}^n \to \mathbb{R}$, f is differentiable at x if

$$\exists q \in \mathbb{R}^n \text{ s.t. } \lim_{h \to 0} \frac{f(x+h) - f(x) - \langle q, h \rangle}{\|h\|} = 0$$

In $\mathbb{R}$, this is the same as saying that the derivative exists, but in higher orders its even more powerful

Note that

$$f \in C^1 \implies f \text{ diff } \implies \text{ all } \frac{\partial f}{\partial x_i} \text{ exist}$$

but none of the converses are true.

Recall: If $f \in C^1(\mathbb{R}^n)$, then

$$f(x+h) - f(x) = \langle \nabla f(x + \theta h), h \rangle \text{ for some } 0 < \theta < 1$$

and as a consequence we proved the above limit.

i.e.

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + E(x, h) = f(x) + \langle \nabla f(x), h \rangle + o(\|h\|)$$

when

$$\lim_{h \to 0} \frac{E(x, h)}{\|h\|} = 0 \text{ or}$$
$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + o(\|y - x\|)$$

for x fixed, y close to x. This has a connection with the tangent plane to the graph of f in two dimensions.

Suppose $G = \{(y_1, y_2, f(y_1, y_2))\}$, and we would like to write down the tangent direction at a point (x_1, x_2). Let there be a curve $\gamma_1(t)$. This is meant to be a parametrized curve. Suppose $\gamma_1(t) = (x_1 + t, x_2, f(x_1 + t, x_2))$. Then the first direction tangent $T_1 = \gamma_1'(0) = (1, 0, \frac{\partial f}{\partial x_1}(x_1, x_2))$.

Similarly let $\gamma_2(t) = (x_1, x_2+t, f(x_1, x_2+t))$. Then the second direction tangent $T_2 = \gamma_2'(0) = (0, 1, \frac{\partial f}{\partial x_2}(x_1, x_2))$.

Then we can let N, the normal vector, be $N = (-\frac{\partial f}{\partial x_1}, -\frac{\partial f}{\partial x_2}, 1)$. Note that N is orthogonal to T_1, T_2.

Then the tangent plane at $(x_1, x_2, f(x_1, x_2))$ is given by

$$\langle y_1, x_1, y_2 - x_2, y_3 - f(x_1 x_2) \rangle \cdot N = 0$$

Note that we are using y to match the former notation. Continuing, we have that

$$y_3 - f(x_1 x_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) - \frac{\partial f}{\partial x_2}(x_1, x_2)(y_2 - x_2) = 0, \text{ or}$$
$$y_3 = f(x_1, x_2) + \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) + \frac{\partial f}{\partial x_2}(x_1, x_2)(y_2 - x_2)$$
$$= f(x_1, x_2) + \langle \nabla f(x_1, x_2), (y_1 - x_1, y_2 - x_2) \rangle$$

which is the above in 2 dimensions. That is to say, we $f(y)$ is the tangent plane approximation to the graph of $f(y)$ at $(x_1, x_2, f(x_1, x_2))$.

8.4

Definition: Let A be an $n \times n$ symmetric matrix. Then,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

We define a quadratic form Q by

$$Q(x) = \langle Ax, x \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

The main application comes from

$$A = \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{bmatrix}$$

for $f \in C^2$. This is called the Hessian matrix.

Proposition: If $f \in C^2(\mathbb{R}^n)$, and with x, h fixed, then

- $\frac{d}{dt} f(x + th) = \langle \nabla f(x + th), h \rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_1}(x + th)h_i$

- $\frac{d^2}{dt^2} f(x + th) = \langle \nabla^2 f(x + th)h, h \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x + th)h_i h_j$

- $\frac{d^3}{dt^3} f(x + th) = \sum_{i,j,k=0}^{n} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x + th)h_i h_j h_k$

Proof:

- For 1, we know from the previous section that

$$f(x + h) - f(x) = \langle \nabla f(x + \theta h), h \rangle$$

  for some $\theta \in (0, 1)$. Then

$$\frac{d}{dt} f(x + th) = \langle \nabla f(x + th), h \rangle$$

- For 2, we build on 1;

$$\frac{d}{dt}[\frac{d}{dt} f(x + th)] = \frac{d}{dt}[\sum_{i=1}^{n} (h_i \frac{\partial f}{\partial x_i}(x + th))]$$

$$= \sum_{i=1}^{n} \frac{d}{dt}[h_i \frac{\partial f}{\partial x_i}(x + th)]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j}(x + th) h_i h_j$$

Definition: Let A be an $n \times n$ symmetric matrix. Then $A = (a_{ij})$ is the notation for the matrix. Define

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

Proposition: aka the Generalized Cauchy Schwarz Inequality. Let A be an $n \times n$ symmetric matrix. Then

$$\|Ax\| \leq \|A\|\|x\| \quad \forall h \in \mathbb{R}^n$$

Proof: We see that

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{h} a_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} x_j \end{bmatrix}$$

, and that

$$\|Ah\|^2 = (\langle \text{ row 1 }, h \rangle^2 + \langle \text{ row 2 }, h \rangle^2 + \dots + \langle \text{ row n }, h \rangle^2)$$

$$= \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} h_j)^2$$

$$\leq (\|a_{i1}\|^2 + \|a_{i2}\|^2 + \dots + \|a_{in}\|^2)(h_1^2 + h_2^2 + \dots + h_n^2)$$

$$= \sum_{i=1}^{n} (\|a_{i1}\|^2 + \|a_{i2}\|^2 + \dots + \|a_{in}\|^2)\|h\|^2$$

$$\leq (\sum_{i,j=1}^{n} \|a_{ij}\|^2)\|h\|^2$$

$$= \|A\|^2\|h\|^2$$

my transcription of this proof is dubious but the idea is there.

Definition: Suppose $A : \mathbb{R}^n \to \mathbb{R}^n$. Then

$$\|A\|_{op} = \sup_{\|x\|=1} \|Ax\|$$

This is the operator norm, which finds the maximum of the norm of the image of the unit sphere.

<u>Remark:</u> The earlier norm definition is also denoted by $\|A\|_{HS}$ and means the Hilbert Schmidt norm. Note that

$$\|Ax\| \leq \|A\|_{HS}$$
$$\|A\|_{op} \leq \|A\|_{HS}$$

Definition: Let A be a symmetric $n \times n$ matrix. Then A is positive definite if $\langle Au, u \rangle > 0 \ \ \forall u \neq 0$, and negative definite if $\langle Au, u \rangle < 0 \ \ \forall u \neq 0$.

<u>Lemma:</u> Let A be a symmetric, positive definite matrix. Then

$$\exists c > 0 \text{ s.t. } \langle Au, u \rangle \geq c\|u\|^2 \ \ \forall u \in \mathbb{R}^n$$

Proof: Think of both sides and think about functions of u. They are both quadratic forms, and homogeneous of degree 2. Basically,

$$\langle Atu, tu \rangle = t^2 \langle Au, u \rangle$$
$$\|tu\|^2 = t^2 \|u\|^2$$

, $\forall t > 0$, $\forall u \in \mathbb{R}^n$. Since they are homogeneous of degree 2, it is thus sufficient to prove that this is true for u on the unit sphere. Then,

$$\langle Au, u \rangle = \langle A \frac{u}{\|u\|}, \frac{u}{\|u\|} \rangle \geq c$$

for some $c > 0$. Basically, the first is true for u iff. it is true for all tu, iff 1 is true for $\frac{u}{\|u\|} = \hat{u}$.

Thus, it suffices to show that $\exists c > 0$ s.t. $\langle A\hat{u}, \hat{u} \rangle \geq c$ for all $\hat{u}$ on the unit sphere. This is true because it's a continuous function on a compact set (the unit sphere), and thus attains a minimum.

The linear algebra proof would have been to diagonalize A and then use the fact that the eigenvalues are positive. Then $c = \min\{\lambda_1, \ldots, \lambda_n\}$.

9 Exam Answers:

Problem 1a: Suppose $f : (0,1) \to \mathbb{R}$ is uniformly continuous. Write down the sequence definition of uniform continuity.

Answer: $\forall \{u_k\}, \{v_k\} \subseteq (0,1)$, if $\{\|u_k - v_k\|\} \to 0$, then $|f(u_k) - f(v_k)| \to 0$, or $\forall \epsilon > 0, \exists \delta > 0$ s.t. $|u - v| < \delta \implies |f(u) - f(v)| < \epsilon$.

Problem 1b: Negate the definition of uniform continuity.

Answer: $\exists \{u_k\}, \{v_k\} \subseteq (0,1)$ such that $\{\|u_k - v_k\|\} \to 0$ but $|f(u_k) - f(v_k)| \not\to 0$, or $\exists \epsilon > 0$ s.t. $\forall \delta > 0, \exists u, v$ with $|u - v| < \delta$ but $|f(u) - f(v)| \geq \epsilon$.

Problem 1c: Is $f(x) = \frac{1}{x}$ uniformly continuous on $(0, 1)$?

Answer: No, because f is not bounded on $(0, 1)$. You can also use $u_k = \frac{1}{k}, v_k = \frac{1}{k^2}$ and then show that the limit of the difference of functions is not 0.

Problem 2: The maximum value of

$$\frac{x_1 - x_2 + \cdots + (-1)^{n+1}x_n}{\sqrt{x_1^2 + \cdots + x_n^2}}$$

Answer: Recognize that this is a C-S problem. Then

$$\frac{\langle (x_1, \ldots, x_n), (1, -1, 1, \ldots, \pm 1) \rangle}{\|x\|} \leq \sqrt{n}$$

so the maximum is $\sqrt{n}$, and the maximizer is $(1, -1, 1, \ldots, \pm 1)$.

Problem 3: Let $\|u\| < 1$. Then $\|u - v\| < 1 - \|u\|$. Prove that $\|u\| < 1$

Answer: $\|v\| \leq \|v - u + u\| \leq \|v - u\| + \|u\| < 1 - \|u\| + \|u\| = 1$.

Problem 4a: Let $f : \mathbb{R}^n \to \mathbb{R}$, continuous. Then prove that

$$\forall V \subseteq \mathbb{R}^n, \text{ open }, f^{-1}(V) \text{ open}$$

Answer: Let $u \in f^{-1}(v)$. Then $f(u) \in V$. For V open, $\exists \epsilon > 0$ s.t. $B_\epsilon(f(u)) \subseteq V$. Since f is continuous, $\exists \delta > 0$ s.t. $B_\delta(u) \subseteq f^{-1}(B_\epsilon(f(u))) \subseteq f^{-1}(V)$. Alternatively $\exists \delta > 0$ s.t. $f(B_\delta(u)) \subseteq B_\epsilon(f(u)) \subseteq V$. Thus $f^{-1}(V)$ is open.

Problem 5a: Let $l > 0$. Then let

$$X_l = \{f : [0, l] \to [1, 3]\}, \text{ continuous}$$

and let

$$T(f)(x) = 1 + \int_0^x f(t)^2 dt$$

Show that $T : X_l \to X_l$.

Answer: We must show that if $f \in X_l$, then $T(f) \in X_l$. Then

$$1 \leq 1 + \int_0^x f(t)^2 dt \leq 3$$

The first part is clearly true, we need that

$$\int_0^l f(t)^2 dt \leq 2 \quad \forall x \in [0, l]$$

An easy bound is to say that

$$\int_0^l f(t)^2 dt \leq l * 3^2 = 9l < 2$$

so $l = \frac{2}{9}$.

Problem 5b: Find a sufficient condition on l such that T is a contraction. That is, find l s.t.

$$\|T(f_1) - T(f_2)\| \leq c\|f_1 - f_2\|$$

Answer: Then

$$T(f_1)(x) - T(f_2)(x) = \int_0^x f_1(t)^2 - f_2(t)^2 dt$$

We want the max of the absolute value of this to be less than or equal to the maximum distance between f_1 and f_2. Then

$$\max_{x \in [0,l]} |T(f_1) - T(f_2)| \leq \max_{x \in [0,l]} \int_0^x |f_1(t)^2 - f_2(t)^2| dt$$

$$\leq \int_0^l |f_1(t) + f_2(t)||f_1(t) - f_2(t)| dt$$

$$\leq \int_0^l 6 \cdot |f_1(t) - f_2(t)| dt$$

so $l \leq \frac{1}{6}$.

Problem 6a: Let $\{f_k\}$ such that $f_k : [-1, 1] \to \mathbb{R}$. Then let $f_k = (1 + x)^k (1 - x)^k$. Prove that

$$\{f_k\} \to 0 \text{ uniformly}$$

Answer: $(1 + x)^k (1 - x)^k = (1 - x^2)^k \leq 1$. Note it goes to 1 if $x = 0$. Otherwise it goes to 0.

Problem 6b: Is f_k cauchy?

Answer: No, because the limit is not continuous.

10 Continuing Differentiating Multivar Functions

10.1 Second Order Approx Formula

<u>Recall:</u> If $f : \mathbb{R} \to \mathbb{R}$, $f''(x)$ exists for every x, then $\quad \forall x, h \in \mathbb{R}$, we have the following:

$$f(x + h) = f(x) + f'(x)h + \frac{1}{2}f''(x + \theta h)h^2 \text{ for some } \theta \in (0, 1)$$

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}, C^2$, and let $x, h \in \mathbb{R}^n$. Then

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2}\langle h, \nabla^2 f(x + \theta h)h \rangle \text{ for some } \theta \in (0, 1)$$

Proof: Let $\phi(t) = f(x + th)$. Then $\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2}\phi''(\theta)$ for some $\theta \in (0, 1)$. Then notice that

$$\phi'(0) = \frac{d}{dt}|_{t=0} f(x + th) = \langle \nabla f(x), h \rangle$$

$$\phi''(t) = \frac{d^2}{dt^2} f(x + th) = \langle \nabla^2 f(x + th)h, h \rangle$$

<u>Note:</u> This is basically Taylor's approximation for linear case with the remainder quadratic. We can expand this out to higher orders.

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}$, or from some open set in $\mathbb{R}^n$, C^2. Then the limit

$$\lim_{h \to 0} \frac{f(x+h) - [f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2}\langle \nabla^2 f(x)h, h \rangle]}{\|h\|^2} = 0$$

Let's call the quantity on the right in the numerator A.

Proof: Then

$$
\begin{aligned}
\lim_{h \to 0} \frac{f(x+h) - A}{\|h\|^2} &= \lim_{h \to 0} \frac{f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2}\langle h, \nabla^2 f(x+\theta h)h \rangle - A}{\|h\|} \\
&= \lim_{h \to 0} \frac{\langle (\nabla^2 f(x+\theta h) - \nabla^2 f(x))h, h \rangle}{\|h\|} \\
&\leq \lim_{h \to 0} \frac{\frac{1}{2}\|(\nabla^2 f(x+\theta h) - \nabla^2 f(x))h\|\|h\|}{\|h\|^2} \quad \text{by the generalized C-S inequality} \\
&\leq \lim_{h \to 0} \frac{\frac{1}{2}\|\nabla^2 f(x+\theta h) - \nabla^2 f(x)\|\|h\|^2}{\|h\|^2} \to 0 \text{ as } h \to 0
\end{aligned}
$$

<u>Recall:</u> A being an $n \times m$ symmetric matrix is positive definite if $\langle Au, u \rangle > 0 \;\; \forall u \neq 0$. Similarly for negative definite, if $\langle Au, u \rangle < 0 \;\; \forall u \neq 0$.

Let $f : O \to \mathbb{R}$, with O open in $\mathbb{R}$. Let $x \in O$. Then there is a strict local minimizer if

$$\exists \delta > 0 \text{ s.t. } f(x) < f(x+h) \;\; \forall 0 \leq \|h\| < \delta$$

Also, x is a strict local maximizer if

$$\exists \delta > 0 \text{ s.t. } f(x) > f(x+h) \;\; \forall 0 \leq \|h\| < \delta$$

Revision: The $\leq$ should be $<$.

Theorem: For simplicity let $f : \mathbb{R}^n \to \mathbb{R}_I, C^2$. If x is such that the gradient $\nabla f(x) = 0$ and the Hessian $\nabla^2 f(x)$ is positive definite, then x is a strict local minimizer.

Proof: We know that $f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2}\langle \nabla^2 f(x)h, h \rangle + R(h)$ where $R(h)$ is the remainder or error. But it's very small, such that

$$\lim_{h \to 0} \frac{R(h)}{\|h\|} \to 0$$

We are looking for a minimizer; we want to show that $f(x+h) > f(x)$ for h small. First, it's clear that $\langle \nabla f(x), h \rangle = 0$. Then

$$
\begin{aligned}
f(x+h) - f(x) &= \frac{1}{2}\langle \nabla^2 f(x)h, h \rangle + R(h) \\
&\geq c\|h\|^2 + R(h)
\end{aligned}
$$

Since $\lim_{h \to 0} \frac{R(h)}{\|h\|} \to 0$,

$$\exists \delta > 0 \text{ s.t. } |R(h) < \frac{c}{2}\|h\|^2 \;\; \forall \|h\| < \delta$$

i.e. $R(h) \geq -\frac{c}{2}\|h\|^2$ if $\|h\| < \delta$. Thus

$$
\begin{aligned}
f(x+h) - f(x) &\geq \frac{c}{2}\|h\|^2 \;\; \forall \|h\| < \delta \\
&> 0 \text{ if } 0 < \|h\| < \delta
\end{aligned}
$$

Thus x is a strict local minimizer. By a similar argument, if $\nabla^2 f(x)$ is negative definite, then x is a strict local maximizer. $\qquad\square$

Remark: What about the converse? Take $f : \mathbb{R}^n \to \mathbb{R}$, of class C^2. Assume x is a strict local minimizer. Then $\nabla f(x) = 0$. What about the hessian matrix? Does it have to be positive definite? No. An example where the strict local minimum is at 0 but the second derivative is not positive definite is $f(x) = x^4$.

Definition: If $f : \mathbb{R} \to \mathbb{R}, C^2$ has a minimizer at x, then $f'(x) = 0, f''(x) \geq 0$. If $f''(x) \leq 0$, then x is a strict local maximizer.

Definition: A matrix is positive semidefinite if $\langle Ax, x \rangle \geq 0$ $\forall x \in \mathbb{R}^n$. Similarly, negative semidefinite if $\langle Ax, x \rangle \leq 0$ $\forall x \in \mathbb{R}^n$.

Theorem: We will show that the converse of the above theorem is true. Let $f : \mathbb{R}^n \to \mathbb{R}, C^2$. If $\nabla f(x) = 0$ and $\nabla^2 f(x)$ is positive semidefinite, then x is a local minimizer.

Proof: Assume f has a local minimizer at x. Look at $\phi(t) = f(x+th)$. Then $\phi \in C^2$, and ϕ has a maximizer at $t = 0$. Then $\phi'(0) = 0$ and $\phi''(0) \geq 0$. Then

$$\phi'(t) = \langle \nabla f(x + th), h \rangle$$
$$\phi''(t) = \langle \nabla^2 f(x + th)h, h \rangle$$

Then $\phi'(0) = 0$ implies that $\langle \nabla f(x), h \rangle = 0$. Then $\phi''(0) \geq 0$ implies that $\langle \nabla^2 f(x)h, h \rangle \geq 0$. Thus $\nabla^2 f(x)$ is positive semidefinite. $\qquad\square$

IN particular, if $f : \mathbb{R}^n \to \mathbb{R}, C^2$, then this has a local minimizer of x, then all the second derivatives of f at x are geq 0

If $f : \mathbb{R}^n \to \mathbb{R}, C^2$, and this has a local maximizer, then the second derivatives of f at x are less than or equal to 0

Proposition: Let U be open in $\mathbb{R}^n$, and $f : U \to \mathbb{R}, C^2$, assume the laplacian

$$\Delta f(x) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(x) > 0 \quad \forall x \in U$$

Then f has no local maximizer in U.

Proof: Assume by contradiction, x is an interior maximizer. Then $\nabla f(x) = 0$ and $\nabla^2 f(x) \leq 0$. But then $\Delta f(x) \leq 0$, a contradiction. Thus f has no local maximizer in U. $\qquad\square$

Note: This is probably transcribed wrong at some point **Theorem:** Let U be a bounded open set, and $\bar{U} = U \cup \partial U$. where ∂U is the bounded seq compact set U. Let $f \in C^2(U)$, f continuous on $\bar{U}$, satisfying

$$\Delta f(x) \geq 0 \quad \forall x \in U$$

Then

$$\max_{\bar{U}} f = \max_{\partial U} f$$

Proof: First, it's obviously true that $\max_{\bar{U}} f \geq \max_{\partial U} f$ because the maximum on the boundary is a subset of the maximum on the closure. The interesting part is the reverse. Consider $f_\epsilon(x) = f(x) + \epsilon \|x\|^2$. Then the laplacian

$$\Delta f_\epsilon(x) = \Delta f(x) + 2n\epsilon > 0$$

41

Then f_ϵ has no local maximizer in U.

$$\max_{\bar{U}} f \leq \max_{\bar{U}} f_\epsilon \leq \max_{\partial U}(f + \epsilon\|x\|^2) \leq \max_{\partial U} f + \epsilon k, k = \max_{\partial U} \|x\|^2$$

Thus $\max_{\bar{U}} f \leq \max_{\partial U} f$. The same argument holds for the minimum. $\qquad\square$

Remark: Suppose

$$f(x, y) = \frac{x^2 y}{x^2 + y^2},$$
$$\frac{\partial f}{\partial x}(0, 0) = 0$$

Is $f C^1$? i.e. are the partials both cont. at 0?

Using homogeneity,

$$f \text{ is } C^\infty \text{ in } \mathbb{R}^n \setminus (0, 0)$$

is homogeneous of degree 2. Then the derivatives are homogeneous of degree 1. Thus, if $f \in C^1(\mathbb{R}^n \setminus 0)$ and is homogeneous of degree k, the derivatives are homogeneous of degree $k - 1$

$$\frac{\partial}{\partial x_i}[f(tx)] = \frac{\partial}{\partial x_i}[t^k f(x)] t \frac{\partial f}{\partial x_i}(tx) \qquad\qquad = t^k \frac{\partial f}{\partial x_i}(x)$$

Remark: Let $x \in \mathbb{R}^n$. Then let the multi index $\alpha = (\alpha_1, \ldots, \alpha_n), a_i = \{0, 1, \ldots\}$

Definition: $|\alpha| = \alpha_1 + \cdots + \alpha_n$

Definition: $\alpha! = \alpha_1! \alpha_2! \ldots \alpha_n!$

Definition: $x^\alpha = x_1^{\alpha_1} \cdot X_2^{\alpha_2} \ldots$

Definition: $\partial^\alpha f(x) = (\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots x_n^{\alpha_n}} f)(x)$

Proposition:

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha$$

, or the multinomial formula.

Proof: For $n = 2$,

$$(x_1 + x_2)^k = \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} x_1^i x_2^{k-i}$$

Let $\alpha = (i, k - i)$. Then $|\alpha| = k$ so this formula can be rewritten as

$$\sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha$$

Now let the formula be true for $n - 1$. Then

$$((x_1 + \cdots + x_{n-1}) + x_n)^k = \sum_{i=0}^{k} \frac{k!}{i!(k-i)!}(x_1 + \cdots + x_{n-1})^i x_n^{k-i}$$

$$= \sum_{i=0}^{k} \frac{k!}{i!(k-i)!}[\sum_{|\beta|=i} \frac{i!}{\beta!} \tilde{x}^\beta] x_n^{k-i}$$

Define $\alpha = (\beta, k - i)$. We have that $|\beta = i|$ Thus the length $|\alpha| = k$. Thus $\frac{1}{(k-i)!\beta!} = \frac{1}{\alpha!}$ Also, $\tilde{x}^\beta x_n^{k-i} = x^\alpha$.
So the sum total is

$$\sum_{|\alpha|=k} \overset{\frac{k!}{\alpha!}}{x^\alpha}$$

and we are done.

Remark: Let $f \in C^k(\mathbb{R}^n)$, and

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \dots$$

Look at $\phi(t) = f(x + th), \phi : \mathbb{R} \to \mathbb{R}, C^1$. Then we know that

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2}\phi''(0) + \cdots + \frac{1}{(k-1)!}\phi^{(k-1)}(0) + \frac{1}{k!}\phi^{(k)}(0)$$

This is the one dimensional taylor expansion, which we know from other places. We would like to express this differently; to express $\phi^{(k)}(t)$ in terms of the partial derivatives of f.

The textbook has that

$$\phi'(t) = \langle (\nabla f)(x + th), h \rangle = (h \cdot \nabla)(x + th) = h_1 \frac{\partial f}{\partial x_1}(x + th) + \cdots + h_n(\frac{\partial f}{\partial x_n})(x + th)$$

$$= [(h_1 \partial_1 + \cdots + h_n \partial_n)f](x + th)$$

and

$$\phi''(t) = \langle (\nabla^2 f)(x + th)h, h \rangle = ((h_1 \partial_1 + \cdots + h_n \partial_n)^2 f)(x + th)$$

$$\phi^{(k)}(t) = ((h_1 \partial_1 + \cdots + h_n \partial_n)^k f)(x + th) = \sum_{|\alpha|=k} \frac{k!}{\alpha!}(h^\alpha \partial^\alpha f)(x + th)$$

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}, C^k$. Then

$$f(x + h) = \sum_{0}^{k-1} \sum_{|\alpha|=j} \frac{1}{\alpha!} \partial^\alpha f(x)h^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(x + \theta h)h^\alpha$$

$$= \sum_{|\alpha| \leq k-1} \frac{1}{\alpha!} \partial^\alpha f(x)h^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(x + \theta h)$$

Notice the similarity with the one dimensional taylor expansion, where

$$\phi(x + h) = \sum_{i=0}^{k-1} \frac{1}{i!}\phi^i(x)h^i + \frac{1}{k!}\phi^{(k)}(x + \theta h)h^k$$

43

11 Review of Lin Alg

Definition: $T : \mathbb{R}^n \to \mathbb{R}^n$ is linear if $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$

Theorem: If $T : rr^n \to \mathbb{R}^n$ is linear, then $\exists! m \times n$ matrix A such that $T(u) = Au \ \ \forall u \in \mathbb{R}^n$

Proof: Also, uniqueness; if $T(u) = Au$ and $T(u) = Bu$, then $A = B$. You can show that A is the matrix of T by showing that $Ae_i = T(e_i)$.

Existence: Let $e_1, \ldots, e_n$ be the standard basis in $\mathbb{R}^n$. Let $u = (u_1, \ldots, u_n) = u_1 e_1 + \cdots + u_n e_n$. Then

$$T(n) = T(u_1 e_1 + \cdots + u_n e_n) = u_1 T(e_1) + \cdots + u_n T(e_n)$$

$$= (T(e_1), \ldots, T(e_n)) \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

Thus $A = (T(e_1), \ldots, T(e_n))$.

Proposition: Let $\mathbb{R}^n \to_A^T \mathbb{R}^m \to_B^S \mathbb{R}^k$, where T, S are linear transforms. Let A, B matrices such that $T(u) = B(u), S(v) = B(v)$. Then $S \circ T(u) = BAu$

Proof: Compute $(S \circ T)(e_i) = S(T(e_i))$ Then

$$S(T(e_i)) = S(Ae_i) = S(a_{1i} e_1 + \cdots + a_{mi} e_m)$$

$$= S(\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix})$$

$$= a_{1i} S(e_1) + \cdots + a_{mi} S(e_m)$$

$$= \begin{bmatrix} S(e_1) & \ldots & S(e_m) \end{bmatrix} \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

$$= BAe_i$$

Theorem: Let $T : \mathbb{R}^n \to \mathbb{R}^n$, linear, invertible (as a function), iff A is invertible, iff $A \neq 0$.

Theorem: Let A be a $n \times n$ matrix. Then A is invertible ($\exists A^{-1} n \times n$ matrix such that $AA^{-1} = A^{-1}A = I$) iff $\exists c > 0$ s.t. $\|Ax\| \geq c\|x\| \ \ \forall x \in \mathbb{R}^n$

Proof: $\Leftarrow$ Assume $\|Ax\| \geq c\|x\|$. Then the null space of A is $\{0\}$. Then A is injective. Then A is surjective because A is a linear map from $\mathbb{R}^n \to \mathbb{R}^n$. Then A is bijective. Then A is invertible.

<u>Remark:</u> This is analogous to: If $f : S \to S$ is 1-1, then f is onto. (Machedon writes S is onto here but I think he means f)

Conversely, if $\exists A^{-1}$ st $AA^{-1} = A^{-1}A$, then find $c > 0$ such that $\|Ax\| \geq c\|x\|$. Hint: generalized C-S inequality.

Let's say something about x. We can write x as $A^{-1}Au = AA^{-1}u$. Let's use the first one. Then

$$\|Ax\| = \|A^{-1}Ax\| \leq \|A^{-1}\|\|Ax\|$$

Then $\|Ax\| \geq \frac{1}{\|A^{-1}\|}\|u\|$ and we are done.

<u>Remark:</u> If V is a vector space with bases $v_1, \ldots, v_n$, and also $w_1, \ldots, w_n$, then the bases are related by a matrix

$$\begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} w_1 & \ldots & w_n \end{bmatrix} C$$

where

$$C = \begin{bmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \ldots & c_{nn} \end{bmatrix}$$

and $v_i = \sum_{j=1}^n c_{ji} w_j$ If the matrix of T with respect to the basis $v_1, \ldots, v_n$ is A, and $w_1, \ldots, w_n$ is B, then $A = CBC^{-1}$, and $(Tv_1, \ldots, Tv_n) = (w_1, \ldots, w_n)A$. This relation is called the change of basis formula.

This won't be on the exam; **Proof:** $T(v_1), \ldots, T(v_n)$ is a row of vectors. We just said that this is equivalent to

$$\begin{bmatrix} T(v_1) & \ldots & T(v_n) \end{bmatrix} = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}$$

We also have that $(v_1, \ldots, v_n) = (w_1, \ldots, w_n)C$, and by linearity $(T(v_1), \ldots, T(v_n)) = (T(w_1), \ldots, T(w_n))C = (w_1, \ldots, w_n)BC$. Then $(v_1, \ldots, v_n)A = (w_1, \ldots, w_n)CA$. Then $A = C^{-1}BC$.

<u>Remark:</u> Let $F : \mathbb{R}^n \to \mathbb{R}^m$. Assume that all partial derivatives of F exist. Define $DF(x)$ to be

$$DF(x) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \ldots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \ldots & \frac{\partial F_m}{\partial x_n} \end{bmatrix}$$
$$= A$$

where A is the derivative matrix of F. Note that the first row of A is the gradient of F_1, the second row is the gradient of F_2, etc. From now on we will distinguish between row/column vectors; from now on ∇f is a row vector.

Proposition: Let $F : \mathbb{R}^m \to]rr^n$ be C^1. Then

$$F(x + h) - F(x) = \begin{bmatrix} \nabla F_1(x + \theta_1 h) \\ \vdots \\ \nabla F_n(x + \theta_m h) \end{bmatrix} h$$

for some $\theta_i \in (0, 1)$. A problem with using the same theta is that the thetas are not necessarily the same (consider each row of $F(x + h) - F(x)$ and you'll believe it but for one row at a time)

Proof: Apply MVT proof for each F_i.

Theorem: Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be C^1. Then

$$\lim_{h \to 0} \frac{F(x + h) - F(x) - DF(x)h}{\|h\|} = 0$$

You can use a norm in the numerator as well if you want to .

Proof: The ith component of the above quantity is

$$\frac{F_i(x + h) - F_i(x) - \langle \nabla F_i(x), h \rangle}{\|h\|} \to 0$$

.

Theorem: Let $F : \mathbb{R}^n \to \mathbb{R}^m$. Fix an x and assume that $\exists A, m \times n$, such that

$$\lim_{h \to 0} \frac{F(x+h) - F(x) - Ah}{\|h\|} = 0$$

Then all partial derivatives of F exist at x, and $DF(x) = A$.

Proof: Look at the ith ocmponent. Then

$$\lim_{h \to 0} \frac{F_i(x+h) - F(x) - \langle A_i, h \rangle}{\|h\|} = 0$$

exists. In particular, for $h = t \cdot e_j$, with t approaching 0, we get that

$$\lim_{t \to 0} \frac{F_i(x+te_j) - F_i(x) - ta_{ij}}{|t|} = 0$$

We want to get rid of the absolute value. Then

$$\lim_{t \to 0} \frac{F_i(x+te_j) - F_i(x) - ta_{ij}}{t} = 0$$

This holds (and is justified by the absolute value version) because a sequence approaches 0 iff the absolute value of the sequence approaches 0. Thus,

$$\lim_{t \to 0} \frac{F_i(x+te_j) - F_i(x)}{t} = a_{ij}$$

and so the partial derivative exists. Then $DF(x) = A$.

Definition: $F : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at x if $\exists A, m \times n$ matrix such that

$$\lim_{h \to 0} \frac{F_i(x+h) - F_i(x) - Ah}{\|h\|} = 0$$

<u>Remark:</u> $F \in C^1(\mathbb{R}^n) \implies F$ is differentiable $\forall x \in \mathbb{R}^n \implies DF(x)$ exists $\forall x \in \mathbb{R}^n$. These are strict implications (eg. they are not equivalent).

What's an example of a function from $\mathbb{R}^2 \to \mathbb{R}$ for which the derivative matrix exists but is not differentiable in this sense? Consider $F(x, y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$ Then the partials exist at $(0,0)$, but the limit of the difference quotient does not exist.

$$\lim_{h \to 0} \frac{F(h) - F(0) - Ah}{\|h\|} = \lim_{h \to 0} \frac{F(h)}{\|h\|} = \lim_{h \to 0} \frac{h^2}{h^2} = 1$$

Notably F is not continuous at $(0,0)$.

11.1   Something

<u>Example:</u> Let $F : \mathbb{R}^n \to \mathbb{R}^m$, C^1. Then assume $F(0) = 0, DF(0)$ satisfies $\|DF(0) - h\| \geq \|h\|$. Prove that $\exists \delta > 0$ such that $\|F(x)\| \geq \frac{1}{2}\|x\|$ $\forall x \in B(0, \delta)$.

<u>Hint:</u> Use first order approximation formula.

We know that
$$\lim_{h \to 0} \frac{F(h) - F(0) - DF(0)h}{\|h\|} = \lim_{h \to 0} \frac{F(h) - DF(0)h}{\|h\|} = 0$$

Then
$$\frac{\|F(h)\|}{\|h\|} = \frac{\|F(h) - DF(0)h + DF(0)h\|}{\|h\|}$$
$$\geq \frac{\|DF(h)\|}{\|h\|} - \frac{\|F(h) - DF(0)h\|}{\|h\|}$$

Since
$$\lim_{h \to 0} \frac{F(h) - DF(0)h}{\|h\|} = 0$$

, $\exists \delta > 0$ such that $\frac{\|F(h) - DF(0)h\|}{\|h\|} < \frac{1}{2}$ $\forall \|h\| < \delta$. Then
$$\frac{\|F(h)\|}{\|h\|} \geq \frac{\|DF(0)h\|}{\|h\|} - \frac{1}{2}$$
$$\geq \|DF(0)\| - \frac{1}{2}$$

Then $\|DF(0)\| \geq \frac{1}{2}$. Then $\|F(h)\| \geq \frac{1}{2}\|h\|$. Continued the next time!!

Remark: Let $F : \mathbb{R}^n \to \mathbb{R}^n, C^1$. Assume that $F(0) = 0$ and $\exists \delta_1 > 0$ such that $\|F(h)\| \geq \frac{1}{2}\|h\|$ $\forall \|h\| < \delta_1$. Then $\exists \delta_2 > 0$ such that $\|DF(0)h\| \geq \frac{1}{2}\|h\|$ $\forall \|h\| < \delta_2$.

Proof: We know that from the first order approximation formula that
$$\lim_{h \to 0} \frac{\|F(h) - F(0) - DF(0)h\|}{\|h\|} = 0$$

Let δ_2 such that $\|F(h) - DF(0)h\| \geq \frac{1}{2}\|h\|$ $\forall \|h\| < \delta_2$. Then
$$\|DF(0)h\| = \|F(h) - (F(h) - DF(0)h) \geq \|F(h)\| - \|F(h) - DF(0)h\|$$

such that $\|F(h)\| \geq \|h\|, \|F(h) - DF(0)h\| \leq \frac{1}{2}\|h\|$. Then $\|DF(0)h\| \geq \frac{1}{2}\|h\|$.

Remark: Is it true that $\|DF(0)h\| \geq \frac{1}{2}\|h\|$ $\forall h \in \mathbb{R}^n$? We proved that this was true for h in a neighborhood of 0. Suppose this is true for all $\|h\| < \delta_2$. Is it possible for this to hold for all h sufficiently small but not for all h? Both the LHS and RHS are homogeneous of degree 1.

Thus, if the inequality holds for all h in a neighborhood of 0, then it holds for all h. (eg you can just scale by some t to get to the neighborhood of 0)

Remark: Under the same assumptions as the first remark in this series of 3, is it true that $\|DF(0)h\| \geq \|h\|$ $\forall h \in \mathbb{R}^n$?

Recall that the argument in parts a, b, show that $\forall 0 < c < 1, \|DF(0)h\| \geq c\|h\|$. Then $\|DF(0)h\| \geq \|h\|$. So the answer is yes.

Remark: For $h \neq 0$, we have shown that $\|DF(h)\|/\|h\| \geq c$, for all $0 < c < 1$. Now take the limit
$$\lim_{c \to 1} \frac{\|DF(0)h\|}{\|h\|} \geq 1$$

Problem 17: Let $f : \mathbb{R}^2 \to \mathbb{R}, C^2$. Assume $x \in \mathbb{R}^n$ is such that

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h \rangle}{\|h\|^2} = 0$$

Prove that

$$\lim_{h \to 0} \frac{\langle \nabla^2 f(x) h, h \rangle}{\|h\|^2} = 0$$

Proof: We know

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h \rangle - \frac{1}{2} \langle \nabla^2 f(x) h, h \rangle}{\|h\|^2} = 0$$

We know that

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h \rangle}{\|h\|^2} = 0$$

Then

$$\lim_{h \to 0} \frac{\langle \nabla^2 f(x) h, h \rangle}{\|h\|^2} = 0$$

and we are done.

Question: is this quantity 0 without even taking the limit? Yes, because the expression is homogeneous of degree 2 on both sides of the equation.

Linalg question: If A is an $n \times n$ matrix, and $\langle Ah, h \rangle = 0 \ \forall h \in \mathbb{R}^n$, then A is not necessarily always 0. For instance, if

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

then $\langle Ah, h \rangle = 0$ for all h.

Question: If $A = \nabla^2 f(x)$, does it follow that A is identically 0? Yes, because $A = \nabla^2 f(x)$ is symmetric. Then $\langle Ah, h \rangle = \langle h, Ah \rangle = \langle h, \nabla^2 f(x) h \rangle = \langle \nabla^2 f(x) h, h \rangle = 0$.

Question: Prove if A is a symmetric matrix, and $\langle Ah, h \rangle = 0 \ \forall h \in \mathbb{R}^n$, then $A = 0$.

Proof: Let A be symmetric. Then A is diagonalizable. Then $A = PDP^{-1}$, where D is diagonal. Then let h be an eigenvector of A. Then $Ah = \lambda h$, and $\langle Ah, h \rangle = \lambda \langle h, h \rangle = 0$. Then $\lambda = 0$. Then $A = 0$. If the matrix is nonzero/nilpotent, then the matrix is not symmetric.

Problem 18: Let $f : \mathbb{R}^2 \to \mathbb{R}, C^2$. This is a pde; let f have boundary values 0, and let $f(x, y) = 0$ on the boundary of the circle. In other words, if $x^2 + y^2 = 1$. Assume that f satisfies the pde

$$\frac{\partial^2 f}{\partial x^2}(x, y) + \frac{\partial^2 f}{\partial y^2}(x, y) + \frac{\partial f}{\partial y}(x, y) = f \text{ if } x^2 + y^2 < 1$$

We want to show that $f(x, y) = 0$ for all $x^2 + y^2 < 1$.

Proof: Recall that if

$$\frac{\partial^2 f}{\partial x^2}(x, y) + \frac{\partial^2 f}{\partial y^2}(x, y) + \frac{\partial f}{\partial y}(x, y) > 0 \text{ at } (x, y), x^2 + y^2 < 1$$

Then f has no interior maximizers. If f has an interior maximizer, look at the maximizer. If it's strict, then the value of the function at the maximizer is greater than 0. So there's no strict interior maximizer. Then f is identically 0.

12   Chain Rule

Proposition 1: Let $\mathbb{R} \to^\gamma \mathbb{R}^n \to^f \mathbb{R}$ for γ, f, C^1. Then

$$\begin{aligned}(f \circ \gamma)'(t) &= \nabla f(\gamma(t))\gamma't(t) \\ &= \langle \nabla f(\gamma(t)), \gamma'(t) \rangle \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t))\gamma_i'(t)\end{aligned}$$

Proof: Recall that if $x, x + h \in \mathbb{R}^n$, then $\exists \theta \in (0,1)$ such that

$$f(x+h) - f(x) = \langle \nabla f(x + \theta h), h \rangle$$

Let $B = \gamma(t+h), A = \gamma(t)$. Then

$$f(\gamma(t+h)) - f(\gamma(t)) = \langle \nabla f(c), \frac{\gamma(t+h) - \gamma(t)}{h} \rangle$$

for some c on the line from $\gamma(t)$ to $\gamma(t+h)$. Then

$$\begin{aligned}\lim_{h \to 0} \frac{f(\gamma(t+h)) - f(\gamma(t))}{h} &= \lim_{h \to 0} \langle \nabla f(c), \frac{\gamma(t+h) - \gamma(t)}{h} \rangle \\ &= \langle \nabla f(\gamma(t)), \gamma'(t) \rangle\end{aligned}$$

and we are done. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition 2: Let $\mathbb{R}^m \to^G \mathbb{R}^n \to^f \mathbb{R}, f, G \in C^1$. Then

$$\nabla(f \circ G)(x) \text{ exists, and}$$
$$\nabla(f \circ G)(x) = \nabla f(G(x)) \cdot DG(x)$$

Note that the result is an m length row vector, $\nabla f(G(x))$ is an n length row vector, and $DG(x)$ is an $m \times n$ matrix.

Ok im writing the matrix dimensions in the wrong order because CS brain but whatever

Proof: LHS:

$$(\frac{\partial(f \circ G)}{\partial x_1}(x) \dots \frac{\partial(f \circ G)}{\partial x_m}(x))$$

By proposition 1, treating x_i as t and keeping the others fixed,

$$\frac{\partial}{\partial x_i}(f \circ G)(x) = \nabla f(G(x)) \cdot (\frac{\partial G}{\partial x_i}(x))$$

So the LHS is

$$\nabla f(G(x))(\frac{\partial G}{\partial x_1}(x) \dots \frac{\partial G}{\partial x_m}(x)) = \nabla f(G(x)) \cdot DG(x)$$

Proposition 3: Let $\mathbb{R}^m \to^G \mathbb{R}^n \to^F \mathbb{R}^k, F, G \in C^1$. Then

$$D(F \circ G)(x) = DF(G(x)) \cdot DG(x)$$

LHS: $m \times k$ matrix. RHS: $n \times k$ matrix times $m \times n$ matrix.

Proof: The LHS:

$$\begin{bmatrix} \nabla(F_1 \circ G) \\ \vdots \\ \nabla(F_k \circ G) \end{bmatrix}$$

By proposition 2, $\nabla(F_i \circ G)(x) = (\nabla F_i)(G(x)) \cdot DG(x)$. Then

$$\begin{bmatrix} (\nabla F_1)(G(x)) \\ \vdots \\ (\nabla F_k)(G(x)) \end{bmatrix} \cdot DG(x) = DF(G(x)) \cdot DG(x)$$

<u>Remark:</u> If $\mathbb{R} \to^\gamma \mathbb{R}^2 \to^f \mathbb{R}, C^1$, and if f, γ are explicit (eg $f(x,y) = e^{x+y}, \gamma(t) = (cos(t), sin(t))$) then computing $\frac{d}{dt}[f \circ \gamma](t)$ can be done with the one dimensional chain rule. eg

$$\frac{d}{dt}[e^{cos(t)+sin(t)}] = e^{cos(t)+sin(t)} \frac{d}{dt}[cos(t) + sin(t)]$$
$$= e^{cos(t)+sin(t)}(-sin(t) + cos(t))$$

<u>Example:</u> Let $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}, C^1$. Assume

$$y \frac{\partial f}{\partial x}(x,y) - x \frac{\partial f}{\partial y}(x,y) = 0$$
$$f(x,0) = f_0(x) \text{ for } x > 0$$

and find f. This is a pde problem.

We will change coordinates to polar. Define

$$\psi(r,\theta) = f(r\cos\theta, r\sin\theta)$$
$$\frac{\partial \psi}{\partial \theta}(r,\theta) = \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)(-r\sin\theta) + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)(r\cos\theta) = 0$$

Then

$$\psi(r,\theta) = \psi(r,0) = f(r,0) = f_0(r)$$

Thus $f(x,y) = f_0(\sqrt{x^2 + y^2})$.

<u>Example:</u> Let $f \in C^1, r > 0, 0 < \theta < 2\pi$.

$$\psi(r,\theta) = f(r\cos\theta, r\sin\theta)$$

Relate $\nabla\psi(r,\theta)$ to $(\nabla f)(r\cos\theta, r\sin\theta)$.

I was too lazy to write the solution

Problem: 15.3, #5: A hint on the homework problem related to harmonic functions. We are given two functions,

$$(u, v) : \mathbb{R}^2 \to \mathbb{R}^2, u, v \in C^2$$

and a function $w : \mathbb{R}^2 \to \mathbb{R}, C^2$. This function w is harmonic and u, v satisfy CR, the Cauchy-Riemann equations. "By slight abuse of notation", let's call $w(u,v)$ a function of u, v, and $u(x,y), v(x,y)$ functions of x, y. Assume that

$$\frac{\partial^2 w}{\partial u^2}(u,v) + \frac{\partial^2 w}{\partial v^2}(u,v) = 0$$

The Cauchy-Riemann equations are:

$$\frac{\partial u}{\partial x}(x,y) = \frac{\partial v}{\partial y}(x,y)$$

$$\frac{\partial u}{\partial y}(x,y) = -\frac{\partial v}{\partial x}(x,y)$$

Note: these come from saying that $\frac{1}{2}(\frac{d}{dx} + i\frac{d}{dy}) = \frac{d}{dz} = 0$. This is the basis of Math 463, complex analysis

Show that

$$(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})w(u,v)(x,y) = 0$$

, that is, $w(u,v)$ is harmonic. Machedon recommends color coding cancellations on the hw (lol)

<u>Remark:</u> Suppose we have

$$\frac{\partial}{\partial x}(w(u(x,y),v(x,y))) = \frac{\partial w}{\partial u}(u,v)\frac{\partial u}{\partial x} + \frac{\partial w}{\partial v}(u,v)\frac{\partial v}{\partial x}$$

This is just an expression of the chain rule. Then

$$\frac{\partial^2}{\partial x^2}[w(u,v)] = [\frac{\partial^2 w}{\partial u^2}(u,v)\frac{\partial u}{\partial x} + \frac{\partial w}{\partial u\partial v}(u,v)\frac{\partial v}{\partial x}]\frac{\partial u}{\partial x} + \frac{\partial w}{\partial u}(u,v)\frac{\partial^2 u}{\partial x^2} + [...]$$

can also do this wrt y^2. We claim that the sum of these two expressions is 0.

Examples of cancellation:

- $\frac{\partial^2 w}{\partial u^2}(u,v)(\frac{\partial u}{\partial x})^2 + \frac{\partial^2 w}{\partial v^2}(u,v)(\frac{\partial v}{\partial y})^2 = 0$

Lemma: If u,v satisfy the $C-R$ equations, then

$$\frac{\partial u}{\partial x^2} + \frac{\partial u}{\partial y^2} = \frac{\partial v}{\partial x^2} + \frac{\partial v}{\partial y^2} = 0$$

Proof:

$$\frac{\partial u}{\partial x^2} = \frac{\partial}{\partial x}[\frac{\partial v}{\partial y}] \text{ by C-R}$$

$$= \frac{\partial}{\partial y}\frac{\partial v}{\partial x}$$

$$= \frac{\partial}{\partial y}[-\frac{\partial u}{\partial y}]$$

$$= -\frac{\partial u}{\partial y^2}$$

Machedon mentions to make sure that you have 3 cancellations in the HW problem.

13 Chapter 16, whatever that is

Let $f : \mathbb{R} \to \mathbb{R}, C^1$. Let $x_0 \in \mathbb{R}$ such that $f'(x) \neq 0$. Then $\exists U$ neighborhood of x_0 such that f is bijective, and f^{-1} is C^1, and $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$. Let $V = f(U)$

Proof: If we know that $f^{-1} : V \to U$ is C^1, then the formula is immediate

$$f(f^{-1}(y)) = y$$
$$f'(f^{-1}(y))(f^{-1})'(y) = 1$$

WLOG let $f'(x_0) > 0$. Let $U = (x_0 - r, x_0 + r)$ be such that $f'(t) > 0 \ \forall t \in [x_0 - r, x_0 + r]$, i.e. f is strictly increasing on $[x_0 - r, x_0 + r]$. Then f is injective on U. It is onto $V = f(U)$ by the intermediate value theorem. Then f is bijective on U. Then f^{-1} is C^1 on V.

Theorem: Let $F : \mathbb{R}^2 \to \mathbb{R}^2, C^1$, assume (x_0, y_0) st

$$DF(x_0, y_0) \text{ is invertible}$$

Then $\exists U$ neighborhood of (x_0, y_0), V neighborhood of $F(x_0, y_0)$ such that $F : U \to V$ is bijective, $F^{-1} : V \to U$ is C^1, and

$$D(F^{-1})(y) = (DF(F^{-1}(y)))^{-1}$$

Remark: If we know that F^{-1} is C^1, then the formula is immediate.

$$F \circ F^{-1}(y) = y$$
$$DF(F^{-1}(y))(F^{-1})'(y) = I$$

Examples:

$$F(x, y) = (x^2 - y^2, 2xy) \text{ i.e. } F(x + y) = (x + iy)^2$$

The determinant

$$\det \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix} = 4(x^2 + y^2) > 0$$

iff $(x, y) \neq (0, 0)$

Thus if $(x_0, y_0) \neq (0, 0) \exists U$ nbhd of (x_0, y_0), V nbhd of $(x_0^2 - y_0^2, 2x_0 y_0)$ s.t. $F : U \to V$ is bijective What about $(0, 0)$? Does there $\exists U$ nbhd of $(0, 0)$, V nbhd of $(0, 0)$ such that $F : U \to V$ is 1-1?

Recall: Let $F : \mathbb{R}^2 \to \mathbb{R}^2, C^1$. Assume (x_0, y_0) such that $DF(x_0, y_0)$ is invertible. Then $\exists U$ neighborhood of (x_0, y_0), V neighborhood of $F(x_0, y_0)$ such that $F : U \to V$ is bijective, $F^{-1} : V \to U$ is C^1, and

$$D(F^{-1})(y) = (DF(F^{-1}(y)))^{-1}$$

Finishing the example, we have that

$$F(x, y) = (x^2 - y^2, 2xy) \text{ i.e. } F(x + y) = (x + iy)^2$$

The hypothesis; this holds at only $(x_0, y_0) \neq (0, 0)$. However, $F(B_1(0)) = B_1(0)$, and the second assertion (that it's onto) holds.

Proof: Let $x = r \cos \theta, y = r \sin \theta$. Then

$$F(x, y) = (r^2 \cos^2 \theta - r^2 \sin^2 \theta, 2r^2 \cos \theta \sin \theta)$$
$$= r^2 (\cos 2\theta, \sin 2\theta)$$

Then $F(r\cos\theta, r\sin\theta) = r(\cos 2\theta, \sin 2\theta)$. Then $F(B_1(0)) = B_1(0)$.

Example 2: Let

$$F(x,y) = (e^x \cos y, e^x \sin y)$$

Then

$$\det \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix} = e^{2x} > 0$$

Hypothesis: the holds everywhere. Is F 1-1? No, because $F(x,y) = F(x, y+2\pi)$. Is it onto? No, because it doesn't hit $(0,0)$. Or rather it's onto $\mathbb{R}^2 \setminus \{(0,0)\}$.

Example: Let $\phi : \mathbb{R}^2 \to \mathbb{R}^2, C^1$ be

$$F(x,y) = (\phi(x,y), \phi^2(x,y))$$

Then the determinant is

$$\det \begin{bmatrix} \phi_x & \phi_y \\ 2\phi\phi_x & 2\phi\phi_y \end{bmatrix} = 2\phi(\phi_y \phi_x - \phi_x \phi_y) = 0$$

When is this matrix invertible? Never, since the determinant is always 0.

Example: Let $F : \mathbb{R}^2 \to \mathbb{R}^2, C^1, \det(DF(x_0, y_0)) = 0$ at some points. Yet F is 1-1, onto $\mathbb{R}^2$. If

$$F(x,y) = (x^3, y^3)$$

then $DF(x,y) = \begin{bmatrix} 3x^2 & 0 \\ 0 & 3y^2 \end{bmatrix}$, and $\det(DF(x,y)) = 0$ at $(0,0)$. Is F^{-1} C^1? No

Recall: TFAE for an $n \times n$ matrix A:

- A is invertible

- $\exists c > 0$ such that $\|Ax\| \geq c\|x\| \quad \forall x \in \mathbb{R}^n$

Definition: $F : O$ open $\subset \mathbb{R}^n \to \mathbb{R}^n$ is stable if $\exists c > 0$ such that $\|F(x) - F(y)\| \geq c\|x - y\| \quad \forall x, y \in O$.

Remark: F stable $\implies F$ is 1-1, or inejective. If $F(x) = F(y)$, then $\|F(x) - F(y)\| = 0$, so $\|x - y\| = 0$, so $x = y$. What's its relationship with lipschitz? F is stable iff F^{-1} is lipschitz.

Proposition: Let A be an $n \times n$ matrix. Assume $\exists c > 0$ such that $\|Ax\| \geq c\|x\| \quad \forall x \in \mathbb{R}^n$. Let B be an $n \times n$ matrix such that

$$\|A - B\| \leq \frac{c}{2}$$

Then,

$$\|Bx\| \geq \frac{c}{2}\|x\| \quad \forall x \in \mathbb{R}^n$$

Proof: We want to say something about $\|Bx\|$. So

$$\begin{aligned} \|Bx\| &= \|Ax + (B-A)x\| \\ &\geq \|Ax\| - \|B - A\|\|x\| \\ &\geq c\|x\| - \frac{c}{2}\|x\| \end{aligned}$$

Theorem: Let $F : O$ open $\subset \mathbb{R}^n \to \mathbb{R}^n, C^1$, and assume $\exists x^*$ such that

$$DF(x^*) \text{ is invertible}$$

Then $\exists U$ neighborhood of x^* such that

- $DF(x)$ is invertible $\quad \forall x \in U$

- F is stable in U, which implies that F is 1-1

Remark:

- Part 1: Look at the determinant of the derivative matrix, and note that it's nonzero and continuous at x^*. Then it's nonzero in a neighborhood of x^*.

- Part 2: Look at $F : B_r(x^*) \to \mathbb{R}^n$. If $x, y \in B_r(x^*)$ then

$$F(x) - F(y) \begin{bmatrix} \nabla F_1(P_1) \\ \vdots \\ \nabla F_n(P_n) \end{bmatrix} [x - y]$$

    for some $P_1, \ldots, P_n$ on the line from x to y. We know that $DF(x^*)$ is invertible, thus $\exists c > 0$ such that $\|DF(x^*)h\| \geq c\|h\| \quad \forall h \in \mathbb{R}^n$. Let the first matrix in the bsal be matrix B. Let $A = DF(x^*)$. Then $\|A - B\| \leq \frac{c}{2}$. Then $\|B(x - y)\| \geq \frac{c}{2}\|x - y\|$. Then F is stable.

Lemma: Let U open in $\mathbb{R}^n$, $F : U \to \mathbb{R}^n, C^1$. Assume $DF(x)$ is invertible $\quad \forall x \in U$. Let

$$E(x) = \|F(x) - y\|^2$$

If E has an interior minimizer at x^*, then $F(x^*) = y$.

Proof: Suppose $X \to^G F(x) - y \to \|F(x) - y\|^2$, i.e. $E(x) = (\phi \circ G)(x)$ where $z \to^\phi z^2$. Assume x^* is a minimizer of E. Then we know that $\nabla E(x^*) = 0$. Then $\nabla E(x^*) = 0$. Then

$$\begin{aligned} \nabla E(x^*) &= (\nabla \phi)(G(x^*)) \cdot DG(x^*) \\ &= (\nabla \phi)(G(x^*)) \cdot DF(x^*) \\ &= 0 \end{aligned}$$

Thus the vector $\nabla \phi(G(x^*))$ is orthogonal to the image of $DF(x^*)$. Then $\nabla \phi(G(x^*)) = 0$. Then $G(x^*) = 0$. Then $F(x^*) = y$.

Recall: For $F : O$ open $\subset \mathbb{R}^n \to \mathbb{R}^n, C^1$, for $x^* \in O$, $DF(x^*)$ invertible, then $\exists U$ neighborhood of x^* such that $DF(x)$ is invertible $\quad \forall x \in U$, and F is stable in U, which implies that F is 1-1.

Theorem: Assume the above. Then $F(U)$ is open.

Proof: let $y_0 \in F(U)$. Then $\exists x_0 \in U$ s.t. $F(x_0) = y_0. \exists r > 0$ s.t. $B_r(\overline{x_0}) \subseteq U$. Let $S = \{x \in U | \|x - x_0\| = r\}$. We know that $\|F(x) - F(x_0)\| \geq c\|x - x_0\| = cr \quad \forall x \in S$. Look at $\min_x \in B_r(\overline{x_0})\|F(x) - y\|$. Then a minimizer exists. We will rule out a bounding minimizer (at an interior minimizer x, $F(x) = y$).
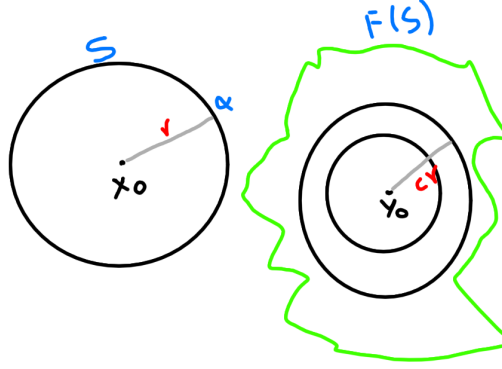
Figure 3: stuf

Let $x \in S$, eg $\|x - x_0\| = r$. Then we know that

$$\|F(x) - y_0\| = \|F(x) - F(x_0)\| \geq cr$$

$$\|F(x) - y\| \geq \|F(x) - y_0\| - \|F(x_0) - y_0\| > \frac{cr}{2}$$

But

$$\|F(x_0) - y\| = \|y_0 - y\| < \frac{cr}{2}$$

and so $x_0 \in B_r(x_0)$, so there's no point in S which can be a minimizer, thus the minimizer must be in the interior of $B_r(x_0)$.

Thus: $\exists U$ nbhd of x^* such that

- $DF(x)$ is invertible $\quad \forall x \in U$

- $\exists c > 0$ such that $\|F(x) - F(y)\| \geq c\|x - y\| \quad \forall x, y \in U$

- $F(U) = V$ is open

- $F^{-1} : V \to U$ is well defined. To show it is C^1 we will prove

$$F(x) = y \text{ exists and equals } (DF^{-1})(y) = (DF(x))^{-1}$$

It suffices to show that

$$\lim_{k \to 0} \frac{\|F^{-1}(y + k) - F^{-1}(y) - [DF(x)]^{-1}k\|}{\|k\|} = 0$$

Notation: $F(x) = y, F(x + h) = y + k$. Then

$$LHS = \frac{\|h - [DF(x)]^{-1}[F(x + h) - F(x)]}{\|k\|}$$

$$= \frac{\|(DF(x))^{-1}[[DF(x)]h - [F(x + h) - F(x)]]\|}{\|k\|}$$

$$\leq \frac{\|DF(x)^{-1}\|\|F(x + h) - F(x) - DF(x)h\|}{\|k\|}$$

We would like to show that $\|k\| \geq C\|h\|$. We know that $\|F(x + h) - F(x)\| = k \geq c\|h\|$. Then we know that the above limit goes to 0 as $k \to 0$ (since $k \to 0 \implies h \to 0$)

<u>Remark:</u> Another proof of inverse thing: Let $F : \mathbb{R}^n \to \mathbb{R}^n, C^1, x^* \in \mathbb{R}^n, DF(x^*)$ invertible. We will show that $\exists \delta_0 > 0$ such that if $\|F(x^*) - y\| < \delta_0$ (divided by some stuff), then $\exists! x \in \bar{B}_{\delta_0}(x^*)$ such that $F(x) = y$.

Proof:

$$F(x) = y \iff x = x - (DF(x^*))^{-1}[F(x) - y]$$

Note that x is obtained as a fixed point as a fixed point of T,

$$x_{k+1} = T(x_k), \{x_k\} \to x$$
$$x_{k+1} = x_k - (DF(x^*))^{-1}[F(x_k) - y]$$

This looks like Newton's method.

diagram here that I can't draw because I don't have my tablet

The main step: To show that $\exists \delta_0$. Then

$$\|x - z - (DF(x^*))^{-1}(F(x) - F(z))\| \leq \frac{1}{2}\|x - z\| \quad \forall x, z \in \bar{B}_{\delta_0}(x^*)$$

The LHS becomes

$$\|(DF(x^*))^{-1}(F(x) - F(z) - DF(x^*(x - z)))\| \leq \|DF(x^*)^{-1}\|\|F(x) - F(z) - DF(x^*(x - z))\|$$

$$= \|DF(x^*)^{-1}\|\| \begin{bmatrix} \nabla F_1(P_1) \\ \vdots \\ \nabla F_n(P_n) \end{bmatrix} - DF(x^*)(x - z)\|$$

by the MVT.

Choose $\delta_0 > 0$ such that the above is less than $\frac{1}{2}$. Then the LHS is $\leq \frac{1}{2}\|x - z\|$. Then T is a contraction, and so T has a unique fixed point.

Next we will show that $T : \bar{B}_{\delta_0}(x^*) \to \bar{B}_{\delta_0}(x^*)$. Let $\|x - x^*\| \leq \delta_0$. Then

$$\|T(x) - x^*\| = \|x - x^* - (DF)(x^*)^{-1}(F(x) - F(x^*) + F(x^*) - y)\|$$
$$\leq \|x - x^* - DF(x^*)^{-1}(F(x) - F(x^*))\| + \|DF(x^*)^{-1}\|\|F(x^*) - y\|$$
$$\leq \frac{1}{2}\|x - x^*\| + \frac{1}{2}\delta_0$$

so T maps $\bar{B}_{\delta_0}(x^*)$ into $\bar{B}_{\delta_0}(x^*)$. Then T has a unique fixed point in $\bar{B}_{\delta_0}(x^*)$, and this fixed point is in $B_{\delta_0}(x^*)$.

<u>Remark:</u> 16.3 #11 problem: Let $F : \mathbb{R}^n \to \mathbb{R}^n, C^1, \exists c$ s.t. $\|F(x) - F(y)\| \geq c\|x - y\|$

c.) To show $F(\mathbb{R}^n)$ is closed, we need $\{y_k\} \subseteq F(\mathbb{R}^n)$, assume $\{y_k\} \to y \in \mathbb{R}^n$. We want $y \in F(\mathbb{R}^n)$. We know that $\exists x_k \in \mathbb{R}^n, F(x_k) = y_k$. If $\{x_k\} \to x$, then $F(x) = y$. You can show that $\{x_k\}$ is cauchy. Then $\{F(x_k)\}$ is cauchy, and so $F(x_k) \to F(x)$.

<u>Remark:</u> Suppose we have $f : \mathbb{R}^2 \to \mathbb{R}, C^1$. When is $\{(x,y)|f(x,y) = 0\}$ a C^1 curve?

A preliminary definition: $C \subseteq \mathbb{R}^2$ is a C^1 curve if $\forall (x_0, y_0) \in C \exists U$ neighborhood of $(x_0, y_0), g : \mathbb{R} \to \mathbb{R}, C^1$ such that $C \cap U = \{(x,y)|y = g(x)\}$.

<u>Fact:</u> $\forall C \subseteq \mathbb{R}^2$, closed, $\exists f :\in C^1(\mathbb{R}^2)$ such that

$$C = \{(x,y)|f(x,y) = 0\}$$

<u>Fact:</u> If $f : \mathbb{R}^2 \to \mathbb{R}$ is C^1, and $\nabla f(x,y) \neq 0$ $\forall (x,y)$ such that $f(x,y) = 0$ then $C = \{(x,y) | f(x,y) = 0\}$ is a C^1 curve.

Theorem: Dim's theorem: Let O open in $\mathbb{R}^2, f : O \to \mathbb{R}, C^1$. Let $(x_0, y_0) \in O$, assume $f(x_0, y_0) = 0, \frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Then

$$\exists, r, R > 0, g : (x_0 - r, x_0 + r) \to (y_0 - R, y_0 + R), C^1 \text{ s.t. } f(x, g(x)) = 0 \;\; \forall |x - x_0| < r$$

additionally, if $(x, y) \in (x_0 - r, x_0 + r) \times (y_0 - R, y_0 + R)$ and $f(x, y) = 0$, then $y = g(x)$.

Proof: $f(x_0, y_0) = 0, \frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. WLOG, $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Then

$$\exists R > 0, c > 0 \text{ s.t. } \frac{\partial f}{\partial y}(x, y) > c \;\; \forall (x, y) \in [x_0 - R, x_0 + R] \times [y_0 - R, y_0 + R]$$

Since $y \to f(x_0, y)$ is strictly increasing on $[y_0 - R, y_0 + R]$, $f(x_0, y_0 - R) < 0, f(x_0, y_0 + R) > 0$. Also, $\exists r > 0$ such that $f(x, y_0 - R) < 0$ if $|x - x_0| < r, f(x, y_0 + R) > 0$ if $|x - x_0| < r$. By the IVT for $y \to f(x, y)$, $\forall |x - x_0| < r \exists y$ s.t. $f(x, y) = 0$. Since $y \to f(x, y)$ is strictly increasing that y is unique. Define $g(x) = y$, the unique y such that $f(x, y) = 0, |y - y_0| < R$. We have constructed $g : (x_0 - r, x_0 + r) \to (y_0 - R, y_0 + R)$ st $f(x, g(x)) = 0$ $\forall |x - x_0| < r$ and if $f(x, y) = 0$ then $y = g(x)$.

To prove $g \in C^1$ first show g is continuous. Let $x, x + h \in (x_0 - r, x_0 + r)$. Then

$$f(x + h, g(x + h)) - f(x, g(x)) = \langle \nabla f(P), (h, g(x + h) - g(x)) \rangle \text{ we get } P \text{ from the MVT}$$
$$= \frac{\partial f}{\partial x}(P)h + \frac{\partial f}{\partial y}(P)(g(x + h) - g(x))$$
$$= 0$$

Then

$$g(x + h) - g(x) = -\frac{\frac{\partial f}{\partial x}(P)}{\frac{\partial f}{\partial y}(P)}h$$

We know that the top is $\leq c$ and the bottom is $\geq c$, so g is C^1 (since $c > 0$)

some stuff missed here about the definition of g'-it's what you would expect so I didn't write it.

<u>Remark:</u> If we change the dimensions of the domain, eg $f : \mathbb{R}^{n+1} \to \mathbb{R}$ of class C^1, and $f(x_0 \in \mathbb{R}^n, y_0) = 0, \frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. By the same proof as above, $\exists r, R > 0, g : (x_0 - r, x_0 + r) \to (y_0 - R, y_0 + R), C^1$ such that $f(x, g(x)) = 0$ $\forall |x - x_0| < r$ and if $f(x, y) = 0$ then $y = g(x)$.

Theorem: The implicit function theorem: We take the notation that $(x, y) \in \mathbb{R}^{n+k}$ where $x \in \mathbb{R}^n, y \in rr^k$. Then let O be the open set in $\mathbb{R}^{n+k}$, $f : O \to \mathbb{R}^k, C^1$. Let $(x_0, y_0) \in O$ such that $F(x_0, y_0) = 0, D_y F(x_0, y_0)$ is invertible. Then

$$\exists R, r > 0, G : B_r(x_0) \to B_R(y_0), C^1 \text{ s.t. } F(x, G(x)) = 0 \;\; \forall x \in B_r(x_0)$$

And additionally; if $x \in B_r(x_0), y \in B_R(y_0)$, and $F(x, y) = 0$, then $y = G(x)$. Also, $DG(x)$ can be computed by the chain rule.

Note that if $k = 1$, then we just have the previous theorem.

Proof: Let $H : O \to \mathbb{R}^{n+k}, H(x, y) = (x, F(x, y))$. Note that $y \in \mathbb{R}^k, x \in \mathbb{R}^n$, so $F(x, y) \in \mathbb{R}^k$. Then,

$$H(x_0, y_0) = (x_0, 0)$$

and the following has a nonzero determinant:

Figure 4: eg $DH(x_0, y_0) = \begin{bmatrix} I & 0 \\ D_x(x_0, y_0) & D_y(x_0, y_0) \end{bmatrix}$

Thus the inverse function theorem applies. Then $\exists R > 0$ and V neighborhood of $(x_0, 0)$ such that $H : B_R(x_0) \times B_R(y_0) \to V$ is 1-1, onto, with a C^1 inverse $H^{-1} : V \to \mathbb{R}^{n+k}, H^{-1}(x, y) = (M(x, y), N(x, y))$.
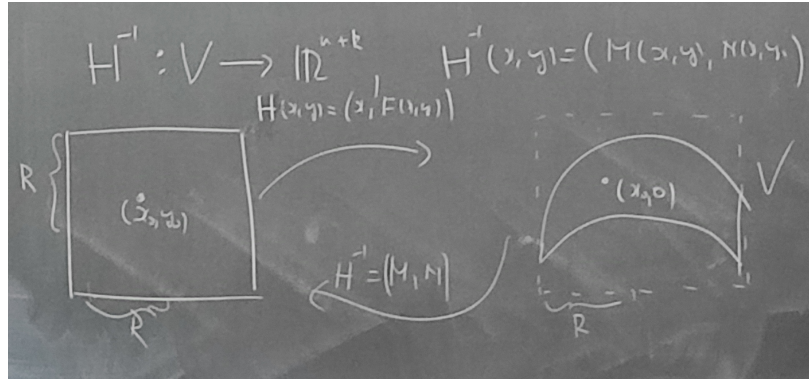

Figure 5: "draw a picture" he said

Use $H(M(x, y), N(x, y)) = (x, y) \ \forall (x, y) \in V, (M(x, y), F(M(x, y), N(x, y))) = (x, y)$, i.e. $(x, F(x, N(x, y))) = (x, y)$. Then $G(x) = N(x, 0)$ is the desired function. Pick $0 < r < R$ such that $B_r(x_0) \times \{0\} \subseteq V$. Notice that $G \in C^1$ and that if $x \in B_r(x_0), y \in B_R(y_0)$, and $F(x, y) = 0$, then we can write $H^{-1}H(x, y) = 0$. Then

$$(M(x, F(x, y)), N(x, F(x, y))) = (x, y)$$

If $F(x, y) = 0$, then $N(x, F(x, y)) = y$, so $y = G(x)$.

Finally, we will write down a formula for $DG(x)$ where $x \in B_r(x_0)$. We use a defining property that $F(x, G(x)) = 0$. Note that a comma denotes a column vector joining; eg

$$x \to (x, G(x)) = \begin{bmatrix} x \\ G(x) \end{bmatrix} \to F(x, G(x)) = 0$$

Then

$$(D_x F(x, G(x)), D_y F(x, G(x))) \begin{bmatrix} I \\ DG(x) \end{bmatrix} = 0$$

$$(D_x F)(x, G(x)) + D_y F(x, G(x)) DG(x) = 0$$

$D_y(x, G(x))$ is invertible in a neighborhood of (x_0, y_0), so we can solve for $DG(x)$. Then

$$DG(x) = -(D_y F(x, G(x)))^{-1} D_x F(x, G(x))$$

Example: Describe a solution for

$$\begin{cases} (x^2 + y^2 + z^2)^3 - x + z = 0 \\ \cos(x^2 + y^2) + e^z - 2 = 0 \end{cases}$$

The LHS: $F(x, y, z), F : \mathbb{R}^3 \to \mathbb{R}^2, F(0, 0) = 0$. We then expect the function to be a curve through $(0, 0, 0) \in \mathbb{R}^3$.

$$DF(0) = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & \end{bmatrix}$$

Then we have that

$$y = (x, z)$$
$$x = y$$
$$F(x, G(x)) = 0$$
$$F(y, G(y)) = 0, G : \mathbb{R} \to \mathbb{R}^2$$

The solutions will then look like $(g_1(y), y, g_2(y))$.

14 Chapter 17

Let $f : \mathbb{R}^3 \to \mathbb{R}, C^1$. Then let $S = \{(x, y, z) | f(x, y, z) = 0\}$, assume $\nabla f(x, y, z) \neq 0 \quad \forall x, y, z \in S$. Then S is a C^1 surface. (Recall a definition for $S \subseteq \mathbb{R}^3$ to be a surface:)

$$\forall x \in S \exists W \text{ nbhd of } x \text{ s.t. } W \cap S = \{(x, y(x)) | y(x) \in \mathbb{R}^2\}$$

Proof: Let $(x_0, y_0, z_0) \in S$. WLOG, $\frac{\partial f}{\partial z}(x_0, y_0, z_0) \neq 0$. By the implicit function theorem $\exists r, R > 0, g : B_r(x_0) \to B_R(y_0), C^1$ such that $f(x, y, g(x, y)) = 0 \quad \forall x, y \in B_r(x_0)$, and these are the only solutions for $f(x, y, z) = 0$ in $B_r(x_0, y_0) \times (z_0 - R, z_0 + R)$.

Theorem: $(x, y) \to (x, y, g(x, y))$ parameterize S near (x_0, y_0, z_0). Let the tangent vectors at (x_0, y_0, z_0):

$$T_1 : (1, 0, \frac{\partial g}{\partial x}(x_0, y_0))$$
$$T_2 : (0, 1, \frac{\partial g}{\partial y}(x_0, y_0))$$

for $x \to (x, y_0, g(x, y_0))$. Then we claim $\nabla f(x_0, y_0, z_0) \perp T_1, T_2$.

Proof: We want to show the dot product is 0. Look at $f(x, y, g(x, y)) = 0 \quad \forall (x, y) \in B_r(x_0, y_0)$. Then

$$0 = \frac{\partial}{\partial x}|_{x=x_0} f(x, y_0, g(x, y_0))$$
$$= \frac{\partial f}{\partial x}(x_0, y_0, z_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0) \frac{\partial g}{\partial x}(x_0, y_0)$$
$$= 0$$

Thus $\nabla f(x_0, y_0, z_0) \perp T_1, T_2$. Thus $\exists \lambda \neq 0$ such that $\nabla f(x_0, y_0, z_0) = \lambda(T_1 \times T_2)$. Then S is a C^1 surface.

<u>Remark:</u> Curves in $\mathbb{R}^3$ defined by the intersection of 2 surfaces: Let $g, h : \mathbb{R}^3 \to \mathbb{R}, C^1$. Let $C = \{(x, y, z)|g(x, y, z) = h(x, y, z) = 0\}$. A sufficient condition for C to be a 1 dimensional curve in $\mathbb{R}^3$ is $\nabla g(x_0, y_0, z_0) \times \nabla h(x_0, y_0, z_0) \neq 0$.

Equivalently, let $G : \mathbb{R}^3 \to \mathbb{R}^2, G(x, y, z) = (g(x, y, z), h(x, y, z))$. We require that

$$DG(x_0, y_0, z_0) = \begin{bmatrix} \nabla g(x_0, y_0, z_0) \\ \nabla h(x_0, y_0, z_0) \end{bmatrix}$$

has rank 2 $\forall (x_0, y_0, z_0) \in C$. If so, WLOG, $D_{y,z}G(x_0, y_0, z_0)$ is invertible. Then by the implicit function theorem, $\exists r, R > 0, \gamma : B_r(x_0) \to B_R(y_0), C^1$ such that $G(x, \gamma(x)) = 0$ $\forall |x - x_0| < r$, a unique solution. Then C agrees with the curve $(x, \gamma(x))$.

Parameterize $C : x \to (x, \gamma(x))$. Take the tangent vector at $(x_0, y_0, z_0) : (1, \gamma'(x_0))$. Then the cross product of the gradients is a nonzero tangent vector, so $\exists \lambda \neq 0$ such that $\nabla g(x_0, y_0, z_0) \times \nabla h(x_0, y_0, z_0) = \lambda(1, \gamma'(x_0))$.

A generalization: For an n dimensional manifold embedded in $\mathbb{R}^N, N = n + k$, we have $F : \mathbb{R}^{n+k} \to \mathbb{R}^k, C^1$. Assume

$$DF(x_0) = \begin{bmatrix} \nabla F_1(x_0) \\ \vdots \\ \nabla F_k(x_0) \end{bmatrix}$$

has maximal rank k. If so, we will represent the level set $M = \{x|F(x) = 0\}$ locally as a graph. Let $X = (x_0, y_0) \in M, x_0 \in \mathbb{R}^n, y_0 \in \mathbb{R}^k$. WLOG $D_y F(x_0, y_0)$ is invertible. Thus $\exists r, R > 0, G : B_r(x_0) \to B_R(y_0)$ such that $F(x, G(x)) = 0$, and is the only solution in $B_r(x_0) \times B_R(y_0)$. Thus $M \cap (B_r(x_0) \times B_R(y_0)) = \{(x, G(x))\}$. Then M is locally a graph.

n linearly independent tangent vectors at (x_0, y_0):

$$\begin{bmatrix} (1, 0, \ldots \frac{\partial G_1}{\partial x_1}) \\ \vdots \\ (0, \ldots, 1, \frac{\partial G_1}{\partial x_n}) \end{bmatrix}$$

The span of these vectors is the tangent space to M at (x_0, y_0). Then $\nabla F(x_0, y_0) \perp$ the tangent space to M at (x_0, y_0). Then the tangent space to M at (x_0, y_0) is $N(DF(x_0))$.

<u>Remark:</u> Suppose you have the function $F : \mathbb{R}^2 \to \mathbb{R}^3, C^1$. We have $O \subseteq \mathbb{R}^2$, open, what is $F(O)$? If $\frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y) \neq 0$ $\forall (x, y) \in O$, then $F(O)$ is a C^1 (smooth) surface.

Generalize: $F : O \subseteq \mathbb{R}^k$, open, $\to \mathbb{R}^N, C^1$. Assume $DF(x_0)$ has some rank $k, N \geq k$. Denote $F = (F_1 \in \mathbb{R}^k, F_2 \in \mathbb{R}^{N-k})$. WLOG, $DF_1(x_0)$ is invertible. Then $\exists U$ neighborhood of x_0, V neighborhood of $F_1(x_0)$, st $F(U) = \{(y, G(y))|y \in V\}$ for some $G : V \to \mathbb{R}^{N-k}, C^1$. Machedon writes $\mathbb{R}^n$ here but that isn't defined; he must have $N - k = n$.

Proof: $DF_1(x_0)$ is invertible. By the inverse function theorem $\exists U$ a nbhd of x_0, V nbhd of $F_1(x_0)$ such that $F_1 : U \to V$ is 1-1, onto, with a C^1 inverse $F_1^{-1} : V \to U$. Denote points in U by x and y in V. We have that $F(x) = (F_1(x), F_2(x)) = (F_1(F_1^{-1}(y)), F_2(F_1^{-1}(y))) = (y, G(y))$.

<u>Example:</u> Problem 28; Suppose we have $F : \mathbb{R}^3 \to \mathbb{R}^2, C^1, F(0, 0, 0) = (0, 0), DF(0, 0, 0) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, $\exists g, h \in C^1, g, h : (-r, r) \to \mathbb{R}, g(0) = h(0) = 0$ such that

- $F(x, g(x), h(x)) = (0, 0)$ $\forall |x| < r$

- $F(g(y), y, h(y)) = (0, 0)$ $\forall |y| < r$

- $F(g(z), h(z), z) = (0, 0) \quad \forall |z| < r$

Which of these holds for sure? The second one, since in the implicit function theorem we have that $D_y F(0, 0, 0)$ is invertible. The free variable is y.

could the others hold? Is it possible for the first to hold? No. We can't use the implicit function theorem because all that provided us is that the second is true. The problem is the partial derivative with respect to x. Assume by contradiction that the first holds. Then

$$DF(0, 0, 0) \cdot \begin{bmatrix} 1 \\ g'(0) \\ h'(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Why is this a contradiction? Because the second row is $(1, 0, 0)$ so you pull a nonzero term.

Example: Problem 41; Suppose $F : \mathbb{R}^2 \to \mathbb{R}^2, C^1, DF(x)$ is positive definite $\quad \forall x \in \mathbb{R}^2$. Prove that F is 1-1.

Recall: $\exists F : \mathbb{R}^2 \to \mathbb{R}^2, C^1, DF(x)$ invertible $\quad \forall x \in \mathbb{R}^2$. Then F is 1-1.

Proof: Assume $F(x) = F(x + h), x \in \mathbb{R}^2$. We must show that $h = 0$. We have that

$$F(x + h) - F(x) = \begin{bmatrix} \nabla F_1(P_1) \\ \nabla F_2(P_2) \end{bmatrix} h$$

Look at $F(x + th)$. We want to write a function that is $\mathbb{R} \to \mathbb{R}$ so that we can use the 1 dimensional MVT in order to keep $P_1 = P_2$. Consider $\phi(t) = \langle F(x + th), h \rangle$. Then $\phi(1) - \phi(0) = \phi'(\theta)$ for some θ in $(0, 1)$. Then

$$\phi(t) = \langle F(x + th), h \rangle$$
$$= \langle DF(x + th)h, h \rangle$$

We are done because we assumed that $\phi(1) - \phi(0) = 0$, and so $\langle DF(x + \theta h)h, h \rangle > 0$ unless $h = 0$.

Example: Let $F : \mathbb{R}^n \to \mathbb{R}^n, C^1$. Assume $DF(x)$ is invertible, and symmetric. Does F have to be 1-1?

Question: Does there exist $F : \mathbb{R}^n \to \mathbb{R}^n, C^1, DF(x)$ invertible $\quad \forall x \in \mathbb{R}^n, F(\mathbb{R}^n)$ compact?

If a function has all points invertible, then by the inverse function theorem, the image is open. Then it can't be compact.

14.1   Lagrange Multipliers

Case 1: Surfaces in $\mathbb{R}^3$. Let $X = (x, y, z) \in \mathbb{R}^3$, let $g : \mathbb{R}^3 \to \mathbb{R}, C^1$. Let $S = \{(x, y, z) | g(x, y, z) = 0\}$. Assume $\nabla g(x) \neq 0$ if $g(x) = 0$. Let $f : \mathbb{R}^3 \to \mathbb{R}, C^1, x_0$ s.t. $f(x_0) \leq f(x) \quad \forall x \in S$ or $f(x_0) \geq f(x) \quad \forall x \in S$. Then $\exists \lambda \in \mathbb{R}$ such that

$$\nabla f(x_0) = \lambda \nabla g(x_0)$$

This is the Lagrange multiplier.

Proof: WLOG $\frac{\partial g}{\partial z}(x_0, y_0, z_0) \neq 0$. By the implicit function theorem $\exists h : B_r(x_0, y_0) \to \mathbb{R}$ s.t. $\{x, y, h(x, y)\}$ agrees with S in a neighborhood of X_0. Then look at $\phi : B_r(x_0, y_0) \to \mathbb{R}, \phi(x, y) = f(x, y, h(x, y)) = f(H)$ for the purposes of the chain rule; has an unconstrained (internal) minimum or maximum at (x_0, y_0), then $\nabla \phi(x_0, y_0) = 0$. By the chain rule

$$\nabla \phi(x_0, y_0) = \nabla f(x_0, y_0, h(x_0, y_0)) DH(x_0, y_0)$$

$$= \nabla f(X_0) \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial h}{\partial x}(x_0, y_0) & \frac{\partial h}{\partial y}(x_0, y_0) \end{bmatrix}$$

61

The column vectors are the tangent vectors to the surface at (x_0, y_0). Thus $\nabla f(X_0) \perp T_1, T_2$. We know $\nabla g(X_0) \neq 0$ is also normal to S. Then $\nabla f(X_0)$ is a scalar multiple of $\nabla g(X_0)$. We call that scalar λ.

Remark: The same argument can be used for the case of $g : \mathbb{R}^n \to \mathbb{R}, C^1, M = \{X | g(X) = 0\}$, assuming $\nabla g(X) \neq 0$ if $g(X) = 0$. Then $f : \mathbb{R}^n \to \mathbb{R}, C^1, x_0 \in M$ such that $f(x_0) \leq f(x) \; \forall x \in M$ or $f(x_0) \geq f(x) \; \forall x \in M$. Then $\exists \lambda \in \mathbb{R}^n$ such that $\nabla f(x_0) = \lambda \nabla g(x_0)$.

Theorem: $\exists \lambda_1, \lambda_2 \in \mathbb{R}$ such that

$$\nabla f(X_0) = \lambda_1 \nabla g_1(X_0) + \lambda_2 \nabla g_2(X_0)$$

Proof: WLOG, $D_{y,z}(g, h)(x_0, y_0, z_0)$ is invertible. By the implicit function theorem $\exists \gamma : (x_0 - r, x_0 + r) \to \mathbb{R}^2, C^1, (x, \gamma(x))$ agrees with C in a neighborhood of X_0. Let $\phi(x) = f(x, \gamma(x)), \phi : (x_0 - r, x_0 + r) \to \mathbb{R}$. Then ϕ has an unconstrained minimum (or maximum) at $x_0, \phi'(x_0) = 0$. By the chain rule,

$$\nabla f(x_0, \gamma(x_0)) \begin{bmatrix} 1 \\ \gamma'(x_0) \end{bmatrix} = 0$$

Recall that $\nabla g(x_0), \nabla h(x_0)$ are linearly independent. Then the vectors are orthogonal to T where T was the matrix above. Then span of $(\nabla g(x_0), \nabla h(x_0))$ is the span of T and so $\nabla f(x_0, \gamma(x_0))$ is a linear combination of $\nabla g(x_0), \nabla h(x_0)$.

Application: (don't tell Adam) Let A be an $n \times n$ symmetric real matrix. Let $\lambda = \min_{\|x\|=1} \langle Ax, x \rangle$. Then λ is an eigenvalue of A.