

Analysis 3

Yeheli Fomberg

326269651

1 Introduction to Topology

1.1 Norms and Metrics

In this course we are going to analyze functions of the form $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ which requires of course knowing analysis 2 but since we will also be using linear analysis tools this course is in some way a continuation of linear analysis as well. We will start by giving basic topological definitions in the euclidean space \mathbb{R}^d , first we define:

$$\mathbb{R}^d = \left\{ (x_1, x_2, \dots, x_d) \mid \begin{matrix} 1 \leq i \leq d \\ x_i \in \mathbb{R} \end{matrix} \right\}$$

And now we can continue to define some more topological terms:

Definition 1.1. The **Euclidean norm** is defined as:

$$\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^d x_i^2}$$

We can similarly define the L_p norm as:

$$\|x\| = \|x\|_p = \sqrt[p]{\sum_{i=1}^d x_i^p}$$

Which satisfies all properties of the norm

Definition 1.2. The **Euclidean metric** is defined as:

$$d_2(P_1, P_2) = \|P_1 - P_2\|_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Notice that it is induced by the Euclidean norm and similarly we can induce L_p metric using L_p norms.

1.2 Sequences

Up until now we didn't have many problems using subscript for indexes of sequences, but now since we have coordinates we must denote a sequence in another similar way ^{superscript} as such $(x^n)_{n=1}^\infty$. To define convergence:

$$\lim_{n \rightarrow \infty} x^n = x \iff \forall i \left(\lim_{n \rightarrow \infty} x_i^n = x_i \right)$$

Definition 1.3. A sequence $(x^n)_{n=1}^\infty$ is called a **Cauchy sequence** if and only if:

$$\lim_{n, m \rightarrow \infty} \|x^n - x^m\| = 0$$

1.3 Topology

Definition 1.4. A **complete metric space** is a metric space M such that every Cauchy sequence in M converges to some limit in M .

Definition 1.5. An **open set** in a Euclidean space is a subset U such that for any $x \in U$ exists $\varepsilon > 0$ such that any y such that any $y \in B_\varepsilon(x)$ satisfies $y \in U$

Definition 1.6. In a topological space X a space a **neighborhood** of a point $x \in X$ is a subset such that exists an open set U such that $p \in U \subset V$

Definition 1.7. A **close set** E in a Euclidean space X is a subset of X such that:

$$(x^n)_{n=1}^\infty \subseteq E \quad x^n \xrightarrow{n \rightarrow \infty} x \implies x \in E$$

Definition 1.8. A topological space is called **compact** if every open cover of X has a finite subcover. An equivalent definition if the space is Euclidean is that for every sequence $(x^n)_{n=1}^\infty$ has a subsequence $(x^{n_k})_{k=1}^\infty$ that converges to a point x in the space.

Definition 1.9. The **closure** of a topological space X is defined as:

$$\text{Cl}(X) = \left\{ x \mid \exists (x^n) : x^n \xrightarrow{n \rightarrow \infty} x \right\}$$

Definition 1.10. The **interior** of a topological space X is defined as:

$$\text{Int}(X) = \{ x \mid \exists r > 0 : B_r(x) \subseteq X \}$$

Definition 1.11. The **boundary** of a topological space X is defined as:

$$\partial X = \left\{ x \in \mathbb{R}^d \mid \forall r > 0 \exists x \in X \wedge \exists y \in X^c : y, z \in B_r(x) \right\}$$

Definition 1.12. A function is continuous at $x \in X$ if for every $\varepsilon > 0$ exists $\delta > 0$ such that for every $\|x - y\| < \delta$

$$\|f(y) - f(x)\| < \varepsilon$$

And we say that a function is continuous on X if it is continuous for every $x \in X$

Definition 1.13. A function is uniformly continuous on X if for every $\varepsilon > 0$ exists $\delta > 0$ such that for every $\|x - y\| < \delta$

$$\|f(x) - f(y)\| < \varepsilon$$

Remark 1.1. An important equivalent, more general definition for continuity is that if f is a function from A to B then if U is an open set in B implies $f^{-1}(U)$ is an open set we say that f is continuous on A .

Definition 1.14. A **connected** space is a topological space that cannot be represented as the union of two or more disjoint non-empty open subsets.

Definition 1.15. A **path** from $x \in X$ to $y \in X$ is a continuous function from the unit interval $[0, 1]$ to X such that $f(0) = x$ and $f(1) = y$

Definition 1.16. A **path connected** space X is a topological space such that exists a path between any two points in X

2 Practice

Phew! These were a lot of definitions... Now it's time for some practice!

Prove that a continuous function $f: A \rightarrow B$ has a maximum in a compact space

By the completeness axiom for the real numbers we know that the set $f(A)$ has a supremum which we will denote S . By the definition of the supremum it is possible to construct a sequence that converges to it which we shall denote $f(x^n)$. We don't know whether x^n converges or not but we know it has a subsequence that converges so:

$$\lim_{k \rightarrow \infty} x^{n_k} = x \quad \text{and} \quad \lim_{k \rightarrow \infty} f(x^{n_k}) = S$$

Since f is close we get $x \in A$ and since it is continuous $f(x) = S$ which shows that it is continuous and also has a maximum.

Prove that a set E is closed if and only if its complement E^c is open

(\Rightarrow)

Suppose that E is closed, and E^c is not open. Then exists $x \notin E$ such that for all $r > 0$ we get $B_r(x) \cap E \neq \emptyset$ which means we can construct a sequence in E that converges to x but $x \notin E$ in contradiction to the assumption that E is close.

(\Leftarrow)

Suppose that E^c is open and E is not closed, then exists a sequence $(x^n)_{n=1}^\infty$ that converges to some $x \in E^c$ which means that for every $r > 0$ that $B_r(x) \cap E \neq \emptyset$ in contradiction to E^c being open.

3 Differentiability

Let $A \in \mathbb{R}^{m \times n}$. We define the linear map $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ by:

$$T(x) = Ax \quad x \in \mathbb{R}^n$$

Let $T: V \rightarrow W$ be a linear transformation between inner product spaces, we define the operator norm to be:

$$\|T\|_{\text{op}} = \|T\| = \sup_{v \neq 0} \frac{\|Tv\|_W}{\|v\|_V}$$

and T is said to be bounded if $\|T\| < \infty$. An important result to prove is that if T :

$$\|T\|_{\text{op}} \leq \left(\sum_{i,j} a_{ij}^2 \right)^{\frac{1}{2}} < \infty$$

since this gives that:

$$\|T(x) - T(y)\| \leq \|T\|_{\text{op}} \|x - y\|$$

which means that T is continuous and even Lipschitz continuous.

Definition 3.1. An *affine function* is a function of the form:

$$T_{A,b}(x) = Ax + b$$

such that $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$

4 Determining critical points

Suppose that $f: U \rightarrow \mathbb{R}$ such that $U \subset \mathbb{R}^n$ is open and $a \in U$.

Definition 4.1. *If exists $r > 0$ such that for any point $x \in B_r(a)$ we have $f(a) \leq f(x)$ then the point is called a **local weak minimal point**. If the inequality is strict it is a **local strong minimal point**. The same applies for a local weak/strong maximum point.*

Definition 4.2. *If f is differentiable in a and $\nabla f(a) = 0$ then a is called a **critical point** of f .*

Definition 4.3. *If a is a critical point of f and for any $r > 0$ exist $x, y \in B_r(a)$ such that $f(x) < f(a) < f(y)$ then a is called a **saddle point** of f .*

ADD EXAMPLES HERE USING PGFPLOTS

Definition 4.4. *Let $f \in C^2(U)$. We define the **Hessian** of f at a to be the matrix:*

$$H_f = H_f(a) = \left[\frac{\partial^2 f(a)}{\partial x_i \partial x_j} \right]_{i,j=1}^n$$

Notice that by Clairaut's theorem we have that the Hessian is symmetrical and thus it defines a quadratic form as follows:

$$\langle H_f(a)v, v \rangle = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j$$

Using Taylor's theorem and the multi-index notation explained earlier we can express the function as such:

$$f(x+h) = f(x) + \sum_{i=1}^n D_i f(x) h_i + \sum_{|\alpha|=2} \frac{D^\alpha f(x) h^\alpha}{\alpha!} + o(\|h\|^2)$$

We can notice that:

$$|\alpha| = 2 \Rightarrow \alpha_1 + \dots + \alpha_n = 2$$

Which means that either way $\alpha! = 2$, so after some algebraic manipulation:

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle H_f(x)h, h \rangle + o(\|h\|^2)$$

Now if assume that x is a critical point of f , we can see that magic happens. We have $\nabla f(x) = 0$ and then:

$$f(x+h) - f(x) = \langle 0, h \rangle + \frac{1}{2} \langle H_f(x)h, h \rangle + o(\|h\|^2) \approx \frac{1}{2} \langle H_f(x)h, h \rangle$$

Since $H_f(x)$ is symmetrical it is similar to a diagonal matrix which we may denote $A = \text{diag}(\lambda_1, \dots, \lambda_n)$, and under the coordinates system corresponding to the eigenvectors of the Hessian we get that:

$$f(x+h) - f(x) \approx \frac{1}{2} \langle H_f(x)h, h \rangle = \sum_{i=1}^n \lambda_i h_i^2$$

And in case $n = 2$ we have more simply:

$$f(x + (a, b)) - f(x) \approx \lambda_1 a + \lambda_2 b$$

And we can use this formula to determine the type of critical points.

Suppose that $a \in U$ is a critical point of f , then the following hold:

1. If $H_f(a)$ is **positive definite** then a is a **local strong minimum**.
2. If $H_f(a)$ is **negative definite** then a is a **local strong maximum**.
3. If $H_f(a)$ is **indefinite** then a is a **saddle point**.

In other cases we can't tell anything about the point and it might be of any type. ADD PROOF
LATER

5 The inverse function theorem

A set $C \subset \mathbb{R}^n$ is called **convex** if and only if:

$$\forall x, y \in C \quad \forall t \in [0, 1] \quad tx + (1 - t)y = y + t(x - y) \in C$$

Proposition 5.1. *Let $U \subset \mathbb{R}^n$ be convex and open, let $f: U \rightarrow \mathbb{R}^n$ be differentiable. If $f'(x)$ is bounded by M then,*

$$\forall x, y \in U \quad \|f(x) - f(y)\| \leq M\|x - y\|$$

Proof. If $f(x) = f(y)$ then we are trivially finished, otherwise we can denote $v = \frac{f(x) - f(y)}{\|f(x) - f(y)\|}$ and the following functions:

$$\begin{aligned} g: \mathbb{R}^n &\rightarrow \mathbb{R} & g(u) &= \langle v, u \rangle \\ h: [0, 1] &\rightarrow U & h(t) &= y + t(x - y) \end{aligned}$$

Differentiating gives:

$$\begin{aligned} Dg &= v \\ Dh &= x - y \end{aligned}$$

We can now define the function:

$$\varphi = g \circ f \circ h$$

And apply Lagrange's theorem to get that exists $c \in (0, 1)$ such that:

$$\varphi'(c) = \frac{\varphi(1) - \varphi(0)}{1 - 0} = \langle v, f(x) \rangle - \langle v, f(y) \rangle = \|f(x) - f(y)\|$$

On the other hand we have:

$$|\varphi'(c)| = \|g'(f \circ h)(f \circ h)'(h)'\| \leq \|v\| \cdot M \cdot \|x - y\|$$

But we know that $\|v\| = 1$ so we get:

$$\|f(x) - f(y)\| \leq M\|x - y\|$$

As wanted. □

TRY SHOWING THIS STILL STANDS FOR THE $|\cdot|_\infty$ NORM.

Corollary 5.2. *If $f \in C^1$ then f is Lipschitz continuous on any compact subset of its domain.*

Theorem 5.3. *The inverse function theorem - Let $U \subset \mathbb{R}^n$ be open, $f \in C^1(U, \mathbb{R}^n)$, and $a \in U$. If $f'(a)$ is invertible then exist open sets $x \in V \subset U$ and $f(x) \in W$ such that $f: V \rightarrow W$ is a bijection and f^{-1} is also differentiable and:*

$$(f^{-1})'(f(x)) = [f'(x)]^{-1}$$

TEDIOUS PROOF TO BE ADDED LATER.

6 Newton-Raphson method

7 The open transformation theorem

Definition 7.1. Suppose $U \subset \mathbb{R}^n$. A function $f \in C^1(U, \mathbb{R}^m)$ is called **regular** in a point $a \in U$ if $\text{rank} J_f(a) = m$. The function is called regular in U if it is regular for any $a \in U$.

Definition 7.2. A mapping $f: U \rightarrow V$ is called **open** if it send any open set W to an open set. That is for any open set $W \subset U$ then $f(W) \subset V$ is also open.

Theorem 7.1. The open mapping theorem - Let $U \subset \mathbb{R}^n$ be an open set, and let $f \in C^1(U, \mathbb{R}^m)$ be regular in U . Then f is an open mapping.

Proof. In the case of $m = n$ the theorem follows directly from the inverse function theorem. Because U is open, and because we know that for any $a \in U$ that $J_f(a)$ is invertible, we get from the inverse function theorem that exist open sets $a \in V_a \subset U$ and $f(a) \in W_a = f(V_a)$. We then have clearly that:

$$f(U) = \bigcup_{a \in U} W_a$$

Since W_a are open for any $a \in U$ we get that $f(U)$ is open as a union of open sets. For any open set $V \subset U$ we can use this exact proof using $f|_V$ instead.

Otherwise we must have $m < n$ because if $n < m$ the rank of the Jacobian can't be m . Considering our case, the rank of the Jacobian is m which means it has m linearly independent columns. Without lose of generality we assume they are the first m columns and denote:

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} = \left[\frac{\partial f_i}{\partial x_j} \right]_{(i,j)=1}^m$$

We now define a new transformation as such:

$$\begin{aligned} F: U &\rightarrow \mathbb{R}^n \\ F(x) &= (f(x), x_{m+1}, \dots, x_n) \end{aligned}$$

We see that the Jacobian of F is:

$$J_F = \begin{pmatrix} \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} & \star \\ 0 & I_{(n-m) \times (n-m)} \end{pmatrix}$$

We notice that F satisfies the conditions for the inverse function theorem and thus it is an open mapping. We also notice that for every open set $V \subset U$ that:

$$f(V) = \pi(F(V))$$

Where π is the projection $\pi: (x_1, \dots, x_n) \mapsto (x_1, \dots, x_m)$. Since it is clear that π is an open mapping we have that f is also an open mapping as a composition of two open mappings as wanted. \square

8 Constrained optimization

Proposition 8.1. *Let $g \in C^1(U, \mathbb{R}^m)$ be regular, $M = \{x \in U \mid g(x) = 0\}$, and $f \in C(U)$. If $a \in M$ and $f(a) \leq f(x)$ for any $x \in M$ then:*

$$\nabla f(a) \in \text{span}\{\nabla g_1(a), \dots, \nabla g_m(a)\}$$

To actually find the minimum we solve the system of equations:

$$\begin{cases} \nabla f(a) = \sum_{i=1}^m \lambda_i \nabla g_i(a) \\ g(a) = 0 \end{cases}$$

Notice that the first equation gives n equations in $n+m$ variables that are $a_1, \dots, a_n, \lambda_1, \dots, \lambda_m$ and the second equation gives m equations in n variables, thus we have $n+m$ equation in $n+m$ variables.

Remark 8.1. *The variables λ_i for $1 \leq i \leq m$ are called **Lagrange multipliers**.*

Here's a simple example to how we can use Lagrange's multiplies to find the distance of a plane from the origin. The function we want to minimize is:

$$f(x, y, z) = \|(x, y, z)\|_2 = \sqrt{x^2 + y^2 + z^2}$$

But this is actually equivalent to finiding the minimum of:

$$f(x, y, z) = x^2 + y^2 + z^2$$

Under the constraints of a plane:

$$g(x, y, z) = ax + by + cz - d = 0$$

We notice that g is indeed regular. To find the minimum we will solve the system of equations:

$$\begin{cases} \nabla f(x, y, z) = \sum_{i=1}^m \lambda_i \nabla g_i(x, y, z) \\ g(x, y, z) = 0 \end{cases}$$

Since g is a function to \mathbb{R} we only have one Lagrange multiplier. We can calculate the gradients of the functions:

$$\nabla f(x, y, z) = (2x, 2y, 2z)$$

$$\nabla g(x, y, z) = (a, b, c)$$

And put them in the system to get:

$$(2x, 2y, 2z) = \lambda(a, b, c) \quad \text{thus} \quad (x, y, z) = \frac{\lambda}{2}(a, b, c)$$

from the first equation. Putting that in the second equation gives:

$$g\left(\frac{\lambda}{2}(a, b, c)\right) = \frac{\lambda}{2}(a^2 + b^2 + c^2) - d = 0$$

Finally we get that the Lagrange multiplier is:

$$\lambda = \frac{2d}{a^2 + b^2 + c^2}$$

And the minimum point is:

$$(x, y, z) = \frac{d}{a^2 + b^2 + c^2}(a, b, c)$$

And the minimal distance of the plane from the origin is:

$$\|(x, y, z)\| = \left\| \frac{d}{a^2 + b^2 + c^2}(a, b, c) \right\| = \frac{d}{\|(a, b, c)\|}$$

In fact what we have shown so far is not a complete proof because the theorem only implies that at a minimum point the equations hold but it doesn't necessarily mean that the point we found is a minimum point. To complete the proof we can choose an arbitrary point on the plane $P_0 = (x_0, y_0, z_0)$, and denote $R = 2\|P_0\|$, and consider the set:

$$S = \{x \in \mathbb{R}^3 \mid g(x) = 0\} \cap \overline{B_R(0)}$$

This set is clearly compact and since $P_0 \in S$ we know that it's not empty and thus we know that f has a minimum in that set, and moreover, we know a priori that this minimum must be the global minimum of the function. Since we only found one point satisfying the equations given by the theorem it must be that minimum point and now the proof is complete.

Proof. To prove the theorem we will prove the contrapositive. Suppose that:

$$\nabla f(a) \notin \text{span}\{\nabla g_1(a), \dots, \nabla g_m(a)\}$$

We need to show that a is not a minimum point. Define the function:

$$\begin{aligned} F: U &\rightarrow \mathbb{R}^{m+1} \\ F(x) &= (f(x), g(x)) \end{aligned}$$

The Jacobian will be:

$$J_F(a) = \begin{bmatrix} \nabla f(x) \\ \nabla_1(x) \\ \vdots \\ \nabla_m(x) \end{bmatrix}$$

Because we assumed $\nabla f(a)$ is linearly independent from the rest of the gradients at a the rank of the Jacobian at a must be $m+1$ and since the rank function is continuous we know that exists a neighborhood U of a where the Jacobian rank is $m+1$ and thus we can apply the open mapping theorem and get that F is an open mapping in that neighborhood. We know that $F(a) = (f(a), 0)$ is in the open set $F(U)$ and also that for a small enough epsilon that $(f(a) - \varepsilon, 0) \in F(U)$ and thus exists $x \in U$ such that $F(x) = (f(a) - \varepsilon, 0) = (f(x), g(x))$ which means that:

1. $g(x) = 0$
2. $f(x) < f(a)$

It follows that $x \in M$ and $f(x) < f(a)$ and thus a is not a minimum point in M . □

9 The implicit function theorem

In mathematics we have so far talked about explicit function - functions that take the elements of one set and transform them into an element of the other. In contrast, we can consider a new way of looking at functions using implicit functions. First, we will consider implicit equations.

Definition 9.1. An *implicit equation* is a relation of the form $R(x_1, \dots, x_n) = 0$ where R is a multivariable function. A vector $x \in \mathbb{R}^n$ that satisfies $R(x) = 0$ is called a *solution*.

We say that the implicit equation $F(x, y)$ where $(x, y) \in U \subset \mathbb{R}^n \times \mathbb{R}^m$ defines y as an implicit function of x if exist $V_1 \times V_2 \subset U$ such that for any $x \in V_1$ exists a unique $y \in V_2$ such that $F(x, y) = 0$. Intuitively the function maps each $x \in V_1$ to its corresponding unique y .

Theorem 9.1. The implicit function theorem - Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be open, $f \in C^1(U, \mathbb{R}^m)$, (a, b) a solution of $f(x, y) = 0$ and also assume that $\det \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)}|_{(a, b)} \neq 0$. Then exists a neighborhood $(a, b) \in V$ and a function $g \in C^1$ defined around a such that:

$$\forall (x, y) \in V \quad f(x, y) = 0 \iff y = g(x)$$

And $g(a) = b$ and $f(x, g(x)) = 0$ for every x near a .

Theorem 9.2. The derivative of an implicit function - Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be open, $f \in C^1(U, \mathbb{R}^m)$, (a, b) a solution of $f(x, y) = 0$ and also assume that $\det \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)}|_{(a, b)} \neq 0$. Then the derivative of the function g from the implicit function theorem is:

$$g'(a) = -[D_y f(a, b)]^{-1} D_x f(a, b)$$

Notice that this theorem is very useful because it allows us to find $g'(a)$ without knowing what is $g(a)$ itself.

ADD PROOFS LATER.

10 Finding roots of polynomials

So it turns out that given a polynomial $p(x) = \sum_{i=0}^n a_n x^n$ then its roots are continuously dependent on the coefficients.

Definition 10.1. Let $p(x) = \sum_{i=0}^n a_n x^n$ be a polynomial. A root of the polynomial with multiplicity of 1 is called a **simple root**.

We want to show that if x_0 is a simple root of a polynomial $P_a(x) = \sum_{i=0}^n a_n x^n$ such that $a = (a_0, \dots, a_n)$ then there exists a neighborhood V of a in \mathbb{R}^{n+1} such that one of the roots of the polynomial:

$$P_b(x) = \sum_{i=0}^n b_n x^n$$

For any $b \in V$ are given as a function $g \in C^1(\mathbb{R}^{n+1})$.

Proof. Define the function:

$$\begin{aligned} f: \mathbb{R}^{n+2} &\rightarrow \mathbb{R} \\ f(a_0, \dots, a_n, x) &= P_a(x) \end{aligned}$$

We notice that the implicit equation $f(a, x) = 0$ gives all the roots of the polynomial $P_a(x)$. Denote x_0 a simple root of $P_a(x)$ where a is the coefficients vector of the polynomial. We can calculate that:

$$f(a, x_0) = 0$$

And since x_0 is a simple root we also know that:

$$\left. \frac{df}{dx} \right|_{(a, x_0)} = \left. \frac{dP_a}{dx} \right|_{x_0} \neq 0$$

Therefore we satisfied the conditions for the implicit function theorem and can conclude that there exists a neighborhood V of (a, x_0) and a function $g \in C^1$ defined around a such that:

$$\forall (a, x) \in V \quad P_a(x) = 0 \iff x = g(a)$$

□

This theorem is of great importance because when using computers and numerical methods to find roots using approximate coefficients we want to know that the roots we find are a good approximation of the real roots we need to find and this theorem shows exactly that.

11 Manifolds

Definition 11.1. Let $k, n \in \mathbb{N}$ such that $k \leq n$. A subset $M \subset \mathbb{R}^n$ is called a C^1 **manifold of dimension k** if for every $a \in M$ exist open sets $U, V \subset \mathbb{R}^n$ such that $a \in U$ and $V \cap \mathbb{R}^k \times \{0_{n-k}\} \neq \emptyset$ and exists a regular function $f \in C^1(U, V)$ such that:

$$f(M \cap U) = (V \cap \mathbb{R}^k \times \{0_{n-k}\}) = \{x \in V \mid x_{k+1} = \dots = x_n = 0\}$$

Recall that the definition is regular is that in any point in the domain the derivative at that point is a surjective linear transformation - or in other words - the rank of the Jacobian is equal to its row number. Because f is a transformation between two open sets of the same dimension we know that f is regular if and only if its derivative is invertible at any point.

Definition 11.2. A function $f \in C^1$ that is invertible and regular is called a **diffeomorphism**. From the inverse function theorem every diffeomorphism's inverse function is a diffeomorphism in itself.

Remark 11.1. A 1-dimensional manifold is sometimes called a **differential curve** and a 2-dimensional manifold is sometimes called a **differential surface**.

Remark 11.2. The title of this section is indeed "manifolds" but in fact we are only talking about **embedded manifolds**. That is to say, we are not talking about manifolds in the abstract sense at all, only about those who are specifically embedded in \mathbb{R}^n .

Proposition 11.1. The following conditions are equivalent:

1. M is a C^1 manifold of dimension k
2. For every $a \in M$ exists a neighborhood U_a and a regular function $g \in C^1(U, \mathbb{R}^{n-k})$ such that:

$$M \cap U = \{x \in U \mid g(x) = 0\}$$

3. Up to permutation of the variables, for every $a \in M$ exists a neighborhood $a \in V \times W$ such that $V \subset \mathbb{R}^k$ and $W \subset \mathbb{R}^{n-k}$ and exists $h \in C^1(V, \mathbb{R}^{n-k})$ such that:

$$M \cap (V \times W) = \text{graph}(h) = \{(x, h(x)) \mid x \in V\}$$

4. For each $a \in M$ exists a neighborhood U , an open set $V \subset \mathbb{R}^k$ and an injective function $H \in C^1(V, \mathbb{R}^n)$ such that:

- (a) $\text{rank} DH = k$
- (b) $M \cap U = H(V)$
- (c) The function $H^{-1}: H(V) \rightarrow V$ is continuous according to the topology on $H(V)$ induced by \mathbb{R}^n .

Remark 11.3. The function H from 4 is called a **parametrization** of $M \cap U$.

Example 11.1. Let $S^n = \partial \mathbb{B}_{n+1} = \{x \mid \|x\| = 1\} \subset \mathbb{R}^{n+1}$ be the n -dimensional sphere. We will show that this sphere is a n -dimensional manifold.

According to 2

For any $a \in M$ we can choose the open set $U = \mathbb{R}^{n+1} \setminus \{0\}$ such that $x \in U$. Define the function:

$$g: U \rightarrow \mathbb{R}$$

$$g(x) = \left(\sum_{i=1}^{n+1} x_i^2 \right) - 1$$

We see that:

$$\nabla g = 2(x_1, \dots, x_{n+1})$$

So g is regular and continuously differentiable in U . Finally we check that:

$$S^n \cap U = S^n = \{x \in U \mid g(x) = 1\}$$

So indeed this verifies that S^n is an n -dimensional manifold.

According to 3

Let $a \in M$, without loss of generality and allowing permutation of variables we can assume that $a_{n+1} < 0$. Now choose $V = \mathbb{B}_n$ and $W = (-\infty, 0)$. Define:

$$h: V \rightarrow \mathbb{R}$$

$$h(v_1, \dots, v_n) = -\sqrt{1 - \sum_{i=1}^n v_i^2}$$

We can verify that indeed:

$$S^n \cap (\mathbb{B}_n \times (-\infty, 0)) = \{(v, h(v)) \mid v \in V\}$$

According to 4

We will show only the case of $S^n = S^2$. Suppose we want to find a parametrization at the point $(1, 0, 0)$. We define the sets:

$$V = (0, \pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

$$U = \{(x, y, z) \mid x > 0\}$$

And the function:

$$H(\theta, \phi) = (\sin(\theta) \cos(\phi), \sin(\theta) \sin(\phi), \cos(\theta))$$

We can see that H is indeed a continuously differentiable continuous injection. We see that:

$$DH = J_H = \begin{pmatrix} \cos(\theta) \cos(\phi) & -\sin(\theta) \sin(\phi) \\ \cos(\theta) \sin(\phi) & \sin(\theta) \cos(\phi) \\ -\sin(\theta) & 0 \end{pmatrix}$$

In our case $k = 2$, so seeing that the determinant of the upper is $\sin(\theta) \cos(\theta)$ which is 0 only when $\theta = \frac{\pi}{2}$ but then the matrix is:

$$\begin{pmatrix} 0 & \sin(\phi) \\ 0 & \cos(\phi) \\ -1 & 0 \end{pmatrix}$$

Which is a matrix of rank 2 so we conclude that $\text{rank} DH = 2$ for every value pair (θ, ϕ) in V . Conditions $\langle b \rangle$ and $\langle c \rangle$ also hold after some algebraic manipulation.

Remark 11.4. Every open set $U \subset \mathbb{R}^n$ is a manifold of dimension n and every vector space $V \subset \mathbb{R}^n$ is a manifold of dimension $\dim(V)$.

ADD PROOF OF THE EQUIVALENCIES.

Definition 11.3. Let $M \subset \mathbb{R}^n$ be a manifold of dimension k . A function $g: M \rightarrow \mathbb{R}^m$ is called **continuously differentiable** if for every $a \in M$ and every parametrization (H, V) in a we have:

$$g \circ H \in C^1(V, \mathbb{R}^m)$$

SHOW THAT THE DEFINITION IS NOT DEPENDENT ON THE CHOICE OF THE COORDINATE SYSTEM.

SHOW THAT THIS DEFINITION IS EQUIVALENT TO exists a neighborhood $a \in U$ and a function $G \in C^1(U, \mathbb{R}^m)$ such that $G|_{M \cap U} = g|_{M \cap U}$ where g is the C^1 function.

12 The tangent space

Definition 12.1. A continuous function $f: [a, b] \rightarrow \mathbb{R}^n$ is called a **continuous path**. If f is also differentiable we say it is a **continuously differentiable path**. The “1- dimensional” set $f([a, b])$ is called a **curve**.

INTUITION FOR THE TANGENT SPACE

Definition 12.2. For each point $p \in \mathbb{R}^n$ we define the **tangent space** to \mathbb{R}^n in p as:

$$T_p(\mathbb{R}^n) = \{v_p = (p, v) \mid v \in \mathbb{R}^n\}$$

In manifolds we have the following definition

Definition 12.3. Let M be a C^1 manifold of dimension k , and let $p \in M$ be a point on the manifold. The **tangent space** to M in point p is defined as:

$$\left\{ p + \gamma'(t_0) \middle| \begin{array}{l} C^1 \ni \gamma: (a, b) \rightarrow M \\ t_0 \in (a, b) \\ \gamma(t_0) = p \end{array} \right\}$$

And we also denote:

$$T_p(M) = \left\{ \gamma'(t_0) \middle| \begin{array}{l} C^1 \ni \gamma: (a, b) \rightarrow M \\ t_0 \in (a, b) \\ \gamma(t_0) = p \end{array} \right\}$$

Remark 12.1. Let $H: V \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$ be a parametrization of a manifold M around p , and $q \in V$ a point such that $H(q) = p$. Then $T_p(M) = [\text{Im}(DH(q))]_p$.

Example 12.1. Consider S^1 and the parametrization:

$$\begin{aligned} H: \left(\frac{\pi}{2}, \frac{3\pi}{2} \right) \\ H(t) = (\cos(t), \sin(t)) \end{aligned}$$

Choose $p = (-1, 0) = H(\pi)$. We see that:

$$DH(\pi) = \left(\begin{array}{c} -\sin(t) \\ \cos(t) \end{array} \right)_{t=\pi} = \left(\begin{array}{c} 0 \\ -1 \end{array} \right) : \mathbb{R} \rightarrow \mathbb{R}^2$$

We notice that indeed:

$$\begin{aligned} \text{Im}(DH(\pi)) &= \text{Sp} \left\{ \left(\begin{array}{c} 0 \\ -1 \end{array} \right) \right\} \\ \Rightarrow T_p(M) &= \left\{ p + a \left(\begin{array}{c} 0 \\ -1 \end{array} \right) \middle| a \in \mathbb{R} \right\} \end{aligned}$$

MORE DEFINITIONS

Definition 12.4. Let $N \subset \mathbb{R}^3$ be a smooth surface. A vector $0 \neq v \in \mathbb{R}^3$ is said to be **normal** to M at a point $p \in M$ if it is orthogonal to $T_p(M)$.

Remark 12.2. If M is equal to the set of roots of a regular function g in a neighborhood of a then a vector normal to M must be co-linear with $\nabla g(p)$. That is because COMPLETE LATER

13 Manifolds with a boundary

Defining manifolds is very convenient, but many interesting sets we want to analyze are not manifolds like any closed interval or half a sphere. That is why in this section we introduce a new object that is almost like a manifold, but slightly different.

Definition 13.1. *The **half k -dimensional space** is defined as:*

$$\mathcal{H}_k = \{x \in \mathbb{R}^n \mid x_n \geq 0\}$$

Definition 13.2. *A subset $M \subset \mathbb{R}^n$ is called a C^1 **manifold with boundary** of dimension k if for every $a \in M$ exists a neighborhood $a \in U$, an open set $V \subset \mathbb{R}^k$ and an injection $H \in C^1(V, \mathbb{R}^n)$ such that:*

1. $\text{rank} DH = k$.
2. $M \cap U = H(V \cap \mathcal{H}_k)$.
3. $H^{-1}: M \cap U \rightarrow V \cap \mathcal{H}_k$ is continuous with respect to the topology induced by \mathbb{R}^n .

The function H is called a **parametrization** of M at a .

Definition 13.3. *Let $M \subset \mathbb{R}^n$ be a manifold with boundary with dimension k . A point $a \in M$ is called a **boundary point** if $a = H(b)$ where $b \in \partial \mathcal{H}_k$. We denote the set of the boundary points of M as ∂M . The rest of the points are called **inner points** and we denote them as $\text{int}(M)$.*

Remark 13.1. *Watch out! the topological definition of a boundary is different from the definition of a boundary for a manifold!*

For practice here are some exercises:

1. Prove that if $V \subset \mathbb{R}^k$ is an open set such that $\mathcal{H}_k \cap V \neq \emptyset$ and if $H \in C^1(V \cap \mathcal{H}_k, \mathbb{R}^n)$ then exists a function $\tilde{H} \in C^1(V, \mathbb{R}^n)$ such that $\tilde{H}|_{V \cap \mathcal{H}_k} = H$
2. Find explicit formulas for a normal vector and a tangent space. Find examples of manifolds with and without boundaries.

14 Length of a path

Definition 14.1. Let $f: [a, b] \rightarrow \mathbb{R}^n$ be a continuous path, and let $P = (t_0 = a < \dots < t_k = b)$. Define:

$$l(f, P) = \sum_{i=1}^k \|f(t_i) - f(t_{i-1})\|$$

And also define:

$$l(f) = \sup_p l(f, p)$$

If $l(f) < \infty$ we say that f has a **length** of $l(f)$.

Proposition 14.1. If $f \in C^1([a, b], \mathbb{R}^n)$ then the path f has length. Moreover, for every partition P we have $l(f, P) \leq \int_a^b \|f'(t)\| dt$, and $l(f) = \int_a^b \|f'(t)\| dt$.

PROOF LATER

Definition 14.2. A **simple smooth curve** is a set $C \subset \mathbb{R}^n$ that is the image of a path $f \in C^1([a, b], \mathbb{R}^n)$. That satisfies:

1. The path f is an injection on (a, b) .
2. For every $t \in [a, b]$ we have $f'(t) \neq 0$.

Show that if f is a injective path then the curve it defines is a C^1 manifold.

Definition 14.3. A curve is called **directed** if we differentiate between its beginning and end.

Definition 14.4. A curve is called **closed** or a **loop** if it satisfies $f(a) = f(b)$.

Definition 14.5. We define the **length** of a curve to be the length of the path it's defined by.

Prove that definition 14.5 is well defined.

Definition 14.6. A parametrization of a curve g such that $\|g'(s)\| = 1$ is called an **arc length parametrization**.

STUFF

15 Line integrals

ADD SOME MORE PRACTICE FOR ARC LENGTH PARAMETRIZATION

Definition 15.1. Let $C \subset \mathbb{R}^n$ be a smooth simple curve with parametrization $C^1 \ni f: [a, b] \rightarrow \mathbb{R}^n$. If ρ is a continuous function on C then we define the **scalar integral** of ρ on C or **type one line integral** as:

$$\int_C \rho ds = \int_a^b \rho(f(t)) \|f'(t)\| dt$$

Notice that when $\rho(x) = 1$ we get the length of the curve. We can use this type of integral when we want to sum values over that curve. For example, a wire can actually be parameterized using a function that gives us its concentration of mass at any given point and then the integral would give the total mass of the wire.

Definition 15.2. Let $S \subset \mathbb{R}^n$. A function $f: S \rightarrow \cup_{p \in S} T_p(\mathbb{R}^n)$ is called a **vector field** if for all $p \in S$ we have $f(p) \in T_p(\mathbb{R}^n)$.

A similar definition exists for smooth manifolds instead of \mathbb{R}^n but there's no reason to get into it right now.

Definition 15.3. Let $f: [a, b] \rightarrow \mathbb{R}^n$ be a continuously differentiable parametrization of a directed, smooth, simple curve C . Let $F: C \rightarrow \mathbb{R}^n$ be a continuous vector field. The line integral of F over C is called a **type two line integral** and is defined to be:

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \sum_{i=1}^n F_i dx_i := \int_a^b F(f(t)) \cdot f'(t) dt$$

16 Integration

Definition 16.1. A **box** in \mathbb{R}^n is a set $R \in \mathbb{R}^n$ such that:

$$R = \prod_{i=1}^k [a_i, b_i]$$

For real numbers $a_i \leq b_i$. If exists i such that $a_i = b_i$ then we say that the box is **degenerate**.

We define the volume of a box to be:

$$\text{Vol}(A) = \prod_{i=1}^k (b_i - a_i)$$

A partition of an n -dimensional box is similar to the partitions of a 2-dimensional box and thus its definition is omitted. If P is a partition of a box R we denote $B \sim Q$ if B is a subbox of the partition. Since the rest of the basic definitions for a Darboux or Riemann integrals are very similar to the second dimensional case, we will only see the way to denote what's necessary. Let f be a bounded function on a box R . We denote the upper and lower Darboux sums with regards to a partition P of R as such:

$$U(f, P) = \sum_{B \sim R} M_Q \cdot \text{Vol}(Q)$$

Where $M_Q = \sup\{f(x) \mid x \in Q\}$ as usual. We say that f is Darboux integrable if:

$$\inf U(f, P) = \sup L(f, P) = I$$

We denote the integral of f over the box as:

$$I = \int_R f \, dV = \int_R f(x) \, dx = \int_R f$$

MORE THEOREMS FROM ANALYSIS 2 ENSUE.

Definition 16.2. *Fubini's theorem* - Let $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ and suppose all the functions:

$$\begin{aligned} F_1(x_1, \dots, x_{n-1}) &= \int_{a_n}^{b_n} f(x) \, dx_n \\ F_2(x_1, \dots, x_{n-2}) &= \int_{a_{n-1}}^{b_{n-1}} f(x) \, dx_{n-1} \int_{a_n}^{b_n} f(x) \, dx_n \\ &\vdots \\ F_n &\equiv I = \int_{a_1}^{b_1} f(x) \, dx_1 \cdots \int_{a_{n-1}}^{b_{n-1}} f(x) \, dx_{n-1} \int_{a_n}^{b_n} f(x) \, dx_n \end{aligned}$$

Are integrable then $I = \int_R f \, dV$.

To calculate the volume of a space Ω we can calculate the integral $\int_R f \, dV$ for the function $f \equiv 1$. For a simple space in \mathbb{R}^3 we can use Fubini's theorem and get:

$$I = \iiint_{\Omega} f(x, y, z) \, dx \, dy \, dz = \int_a^b dx \int_{y_1(x)}^{y_2(x)} dy \int_{z_1(x)}^{z_2(x)} f(x, y, z) \, dz$$

For the relevant functions.

17 Zero volume

Remark 17.1. If $f: R \rightarrow \mathbb{R}^n$ is a continuous function on R then it is integrable.

Definition 17.1. A set $A \subset \mathbb{R}^n$ is said to have **volume zero** if for each $\varepsilon > 0$ exists a finite number of boxes R_1, \dots, R_k such that:

$$\sum_{i=1}^k \text{Vol}(R_i) < \varepsilon \quad \text{and} \quad A \subset \bigcup_{i=1}^k \text{int}(R_i)$$

We see that any point $x \in \mathbb{R}^n$ is of volume zero because we can always choose one box of arbitrarily small volume such that x is inside it. It is also trivial to prove that a finite union of sets of volume zero is of volumes zero.

Theorem 17.1. Let $f: R \rightarrow \mathbb{R}^n$ be bounded and suppose:

$$X = \{x \in R \mid f \text{ is not continuous in } x\}$$

be a set of volume 0. Then f is integrable over R .

Proof. We know that f is bounded so we can denote its bound M , and also we know that for any $\varepsilon > 0$ exist R_1, \dots, R_k such that:

$$\sum_{i=1}^k \text{Vol}(R_i) < \frac{\varepsilon}{4M} \quad \text{and} \quad X \subset \bigcup_{i=1}^k \text{int}(R_i)$$

Thus we can choose a partition P_X such that for every i we have $R_i \in P$ and then we have:

$$U(f, P) - L(f, P) = \sum_{i=1}^k (M_Q - m_Q) \text{Vol}(R_i) < \frac{\varepsilon}{2} \quad \text{and} \quad X \subset \bigcup_{i=1}^k \text{int}(R_i)$$

Now if we consider the set $Y = R \setminus \{R_i\}_i$ then it is clearly closed and bounded and thus it is integrable and we can choose a partition P_Y of Y such that:

$$U(f, P_Y) - L(f, P_Y) = \sum_{Q \sim P_Y} (M_Q - m_Q) \text{Vol}(Q) < \frac{\varepsilon}{2}$$

And now it is clear that the partition $P = P_X \cup P_Y$ satisfies:

$$U(f, P) - L(f, P) < \varepsilon$$

□

Definition 17.2. We say that $A \subset \mathbb{R}^n$ is of **measure zero** if exist $\{R_i\}_{i \in I}$ such that $|I| < \aleph_0$ and:

$$\sum_{i \in I} \text{Vol}(R_i) < \varepsilon \quad \text{and} \quad A \subset \bigcup_{i \in I} \text{int}(R_i)$$

Remark 17.2. A set of volume zero is always of measure zero, but the contrary is not always true. For example the sets of the form $\mathbb{Q}^n \subset \mathbb{R}^n$ are all of measure zero but not of volume zero.

Theorem 17.2. A compact set $D \subset \mathbb{R}^n$ is of volume zero if and only if it is of measure zero.

The proof of this theorem is very direct and thus omitted.