

Analysis 2 - Thomas Mountford

Alp Ozen

Spring 2019

Contents

1	Notions on \mathbf{R}^n	2
1.1	Introducing topological properties on \mathbf{R}^n	2
1.2	Difficult problems and notes from week 2	4
1.3	More norms on \mathbf{R}^n	4
1.4	Curves in \mathbf{R}^n	5
1.5	Real-valued functions in \mathbf{R}^n	6
1.6	Partial derivatives	6
1.7	Vector Fields	8
1.8	Implicit function theorem	9

1 Notions on \mathbb{R}^n

1.1 Introducing topological properties on \mathbb{R}^n

Let's recall that \mathbb{R}^n is a Euclidean vector space. We define a scalar product on \mathbb{R}^n as follows:

Definition 1.

$$\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

1. $\langle x, x \rangle \geq 0$
2. $\langle x, y \rangle = \langle y, x \rangle$
3. $\langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$

A *norm* is defined as function that maps some real vector space E to \mathbb{R} and satisfies:

- 1) $|x| \geq 0 \forall x \in E, |x| = 0 \iff x = 0$
- 2) $|\lambda \cdot x| = |\lambda| \cdot |x|$
- 3) $|x + y| \leq |x| + |y|$

In our intuitive understanding of \mathbb{R}^n we are actually thinking about the Euclidian space \mathbb{R}^n equipped with the Euclidian norm.

Definition 2. Euclidian norm

$$|x|_2 = \sqrt{\langle x, x \rangle} = \left(\sum_i^n x_k^2 \right)^{\frac{1}{2}}$$

And from this naturally follows the definition of Euclidian distance:

Definition 3.

$$d(x, y) = |x - y|$$

We note that d satisfies the same 3 properties as the norm. Thus, the couple (E, d) is called a metric space.

And now more definitions:

Definition 4. Open sets

1. **Open ball** $B(a, r) := \{x \in \mathbb{R}^n : d(x, a) < r\}$
2. **Open subset** Some subset $S \subset \mathbb{R}^n$ is open if $\forall x \in \mathbb{R}^n, \exists \epsilon > 0 B(x, \epsilon) \subset S$
3. **Closed subset** Some S is closed if $\mathbb{R}^n - S$ is open, note that the empty set and \mathbb{R}^n are both open and closed.
4. **The interior and boundary of a set** a is in the interior of S if $\exists \epsilon > 0 B(a, \epsilon) \subset S$ and b is in the boundary of a set S if any $B(a, \epsilon)$ contains points from both S and $\mathbb{R}^n - S$. The set of all interior points is denoted $(^\circ S)$ and set of all boundary points is denoted ∂S
5. **Closure of a set** a is a closure of S if for any $B(a, \epsilon)$ we have $B(a, \epsilon) \cap S \neq \emptyset$

Definition 5. Topology

A Topology exists whenever the following are satisfied:

$$\text{For a given } M \subset \mathbb{R}^n \text{ we define } O \subset P(M)$$

1. $\emptyset \in O, M \in O$

$$2. U \in O, V \in O \rightarrow U \cap V \in O$$

$$3. U_\alpha \in O \rightarrow \bigcup_\alpha U_\alpha \in O$$

Definition 6. Closure of a set A point $a \in \mathbb{R}^n$ is a closure point of S if for any $B(a, \epsilon)$ we have:

$$B(a, \epsilon) \cap S \neq \emptyset$$

The set of all closure points called the closure of S is denoted \bar{S} noting that this only holds in \mathbb{R}^n

$$\bar{S} = S \cup \partial S$$

Theorem 1.1. Important results on closures and boundaries

$$S^\circ \subset S \subset \bar{S}$$

$$\bar{S} = S^\circ \cup \partial S$$

$$S \text{ is open iff } S = S^\circ$$

$$S \text{ is closed iff } S = \bar{S}$$

Definition 7. Sequence in \mathbb{R}^n

$f : \mathbb{N} \rightarrow \mathbb{R}^n$ such that:

$$x_k = f(k) \in \mathbb{R}^n \forall k \in \mathbb{N}$$

Definition 8. Convergence

The definition of convergence in \mathbb{R}^n is very similar to its counterpart in \mathbb{R} . We say that x_k converges to x iff:

$$\forall \epsilon > 0 \exists N_\epsilon \forall k \geq N_\epsilon d_2(x_k, x) < \epsilon$$

Definition 9. Complete spaces aka. Banach spaces

A space is called complete if every Cauchy sequence in this space converges to a limit. Some example of complete spaces are \mathbb{R}, \mathbb{R}^n

We restate the axioms of a **Norm** and of a **Metric**. The subtle difference between the two is that a norm can only be applied to some vector space whereas a metric is applicable to other spaces.

Remark 1.1. Axioms of norm and metric

Norm:

$$1. \|x\| \geq 0$$

$$2. \|\lambda x\| = |\lambda| \|x\|$$

$$3. \|x + y\| \leq \|x\| + \|y\|$$

Metric:

$$1. d(x, y) \geq 0$$

$$2. d(x, y) = d(y, x)$$

$$3. d(x, y) \leq d(x, z) + d(z, y)$$

1.2 Difficult problems and notes from week 2

We note that any open set is equal to its interior.

Example 1.1. Consider the set defined as $T = \{(x, y) \in \mathbb{R}^2 | 1 < x^2 + 4y^2 < 4\}$ We observe the following about this set:

$$T^\circ = T$$

$$\partial T = \{(x, y) \in \mathbb{R}^2 | 1 = x^2 + 4y^2 \text{ or } x^2 + 4y^2 = 4\}$$

Remark 1.2. Rather interestingly, the rational numbers are neither closed nor open because both the irrationals are dense as well as the rationals being dense which was proved in analysis 1. We have the following properties for \mathbb{Q}

$$\partial \mathbb{Q} = \bar{\mathbb{Q}} = \mathbb{R}$$

Definition 10. If x_k is in a bounded closed set D then \exists a convergent subsequence with limit in D .

Definition 11. Continuity open ball definition

$\forall a \in \mathbb{R}^n$ we have $\forall \epsilon > 0 \exists \delta > 0$ such that $f(B(a, \delta)) \subset B(f(a), \epsilon)$

Definition 12. Continuity open set definition

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ Then f is continuous on \mathbb{R}^n iff. for all $O \subset \mathbb{R}^m$ where O is open, we have $f^{-1}(O)$ open in \mathbb{R}^n

Proof. (only if part) Let f be continuous, we want to show that whenever O is open, $f^{-1}(O)$ is also open. Now take $f^{-1}(O)$. Let $p \in f^{-1}(O)$. Then $f(p) \in O$ Now we have that O is open hence $\exists \epsilon > 0$ s.t. $B(f(p), \epsilon) \subset O$ Now because f is continuous at p we have that $\exists \delta > 0$ s.t. $\underbrace{f(B(p, \delta)) \subseteq B(f(p), \epsilon)}_{\text{applying open ball cont.}} \subseteq O$

This further shows that $B(p, \delta) \subseteq f^{-1}(O)$ which shows that f is open in S .

(if part) Now suppose that for all open sets $O \in \mathbb{R}^m$ we have $f^{-1}(O)$ are open in S . To show that f is continuous, we must show f is continuous on all $p \in S$. Let $p \in S$ s.t. $y = f(p)$. Then $\forall \epsilon > 0$ $B(y, \epsilon)$ is open. By assumption then, $f^{-1}(B(f(p), \epsilon))$ is open in S . Now, $f(p) \in B(f(p), \epsilon)$ we have $p \in f^{-1}(B(f(p), \epsilon))$ is open in S , $\exists \delta > 0$ s.t. $B(p, \delta) \subseteq f^{-1}(B(f(p), \epsilon))$ and applying f to both sides we get our result. \square

1.3 More norms on \mathbb{R}^n

Consider the non-usual norms $|x|_1 := \sum |x|$ and $|x|_\infty = \max |x_k|$. A super interesting result is that all norms on \mathbb{R}^n are equivalent.

Definition 13. Two norms $|\cdot|$ and $\|\cdot\|$ are **equivalent** if $\exists C_1, C_2 > 0$ such that $\forall x \in E$:

$$C_1|x| \leq \|x\| \leq C_2|\cdot|$$

Theorem 1.2. Any two norms on \mathbb{R}^n are equivalent.

We do not prove this theorem but consider the visual below for why it is true.

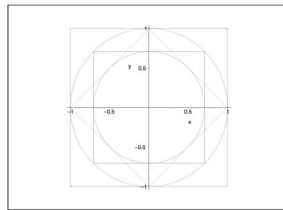


Figure 1: Norm equivalence

1.4 Curves in \mathbb{R}^n

Definition 14. Curve A curve is a function $f : I \rightarrow \mathbb{R}^n$ denoted as:

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

A good example of a curve is the speed of a particle in mechanics. It takes only one input t and has a different function defined for each coordinate. Another example is the graph $G_f = \begin{bmatrix} x \\ f(X) \end{bmatrix}$ of any function from $I \rightarrow \mathbb{R}^k$. Another curve is a helix defined as:

Definition 15. Helix

$$f(t) = \begin{bmatrix} r \cos t \\ r \sin t \\ ct \end{bmatrix}$$

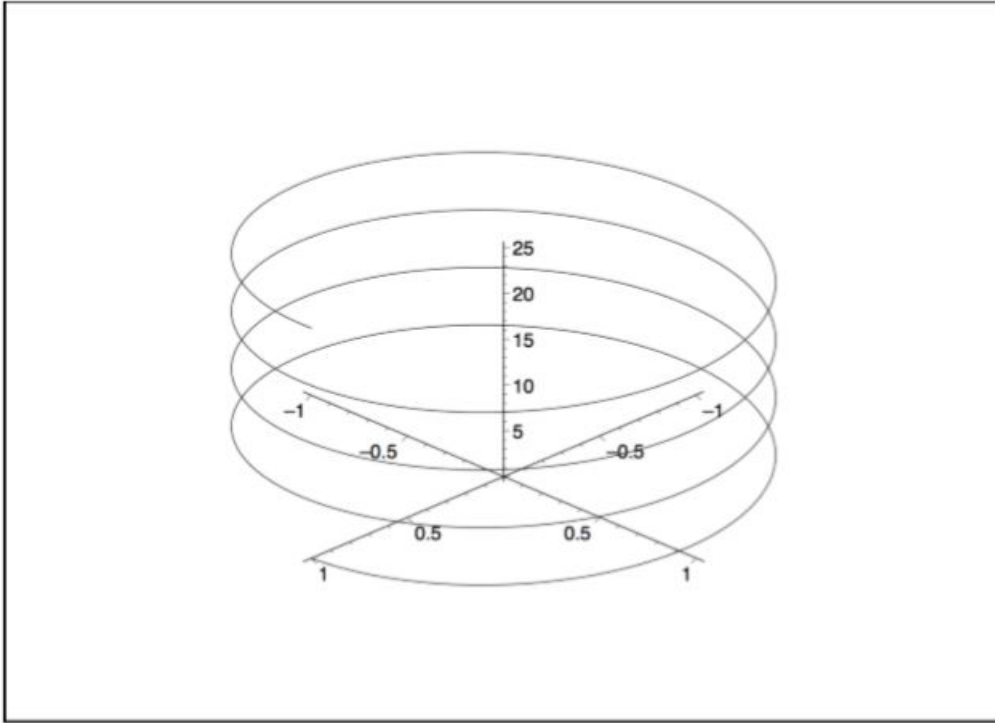


Figure 2: Helix

Definition 16. Arc-length of a curve Given a curve is rectifiable (each segment tends to 0 and sum of segments tend to $L > 0$) we define

$$L = \int_b^a |f'(t)|_2 dt$$

Now the above formula is taught at high school level as:

$$L = \int_b^a \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy$$

which is simply the application of our definition to \mathbb{R}^2

1.5 Real-valued functions in \mathbb{R}^n

Definition 17. *Real valued function* $f : D \rightarrow \mathbb{R}$ such that $D \subset \mathbb{R}^n$

The graph G_f of a real valued function is the set $G_f = \left\{ \begin{bmatrix} x \\ f(x) \end{bmatrix} \mid x \in D \right\} \subset \mathbb{R}^{n+1}$

Hence where a curve is map from one dimension to n dimensions, a real-valued function is a map from n dimensions to 1 dimension.

1.6 Partial derivatives

Definition 18. $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is partially differentiable with respect to variable x_k if:

$$\frac{\partial f}{\partial x_k}(a) = \lim_{h \rightarrow 0} \frac{f(a + he_k) - f(a)}{h}$$

exists. To calculate the partial derivative at x_k we simply let all x_i be constants.

Definition 19. Gradient The gradient of f is defined as:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix} = \sum_{k=1}^n D_k f(x) e_k$$

Example 1.2. Let $f(x) = |x|_2$ Then we have that $\frac{\partial f}{\partial x_k}(a) = \frac{x_k}{r}$

Now we come to a more complicated example:

Example 1.3.

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

In this case, we have that although the partial derivative exists its existence does not imply continuity. The partial derivative itself must be continuous to imply continuity.

Definition 20. Continuity in \mathbb{R}^n

We say that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous at $a \in \mathbb{R}^n$ if $\forall \epsilon > 0 \exists \delta > 0$ such that $|x - a|_2 < \delta \rightarrow |f(x) - f(a)|_2 < \epsilon$

The problem with the above definition is that we may approach say $(0, 0)$ from infinite number of ways (easily seen considering a cartesian plane). This was not the case in $\mathbb{R}^{\mathbb{K}}$ because we could approach 0 from the left or the right. Hence to prove that a limit exists by using cases is impossible, we need to instead use the epsilon-delta definition.

Definition 21. We define $C^1(U)$ as the set of functions $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ such that f has n partial derivatives each continuous.

Definition 22. Contour graphs When we are presented with a real-valued function of multiple variables, we may either represent it via its graph, that is map it to \mathbb{R}^{m+1} or instead draw its contour lines that is we let the function equal some value $c \in \mathbb{R}$ and plot all the points in \mathbb{R}^n that take this value.

And now we define the total differential.

Definition 23. Total differential

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

Well when might the total differential ever be useful? Here's a simple example.

Example 1.4. Let's define $z = \sqrt{x^2 + y^2}$. Now suppose we want to evaluate the function at $(x = 3.92, y = 4.01)$. Then we may simply evaluate the function at $(x = 4, y = 4)$ and using the total differential find the change in z and subtract it from our initial result.

Definition 24. Bilinear form We may think of a bilinear form as a linear map that accepts 2 arguments instead of one and linearity holds for both arguments.

$$B : V \times V \rightarrow F$$

$$B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$$

$$B(fv, w) = fB(v, w)$$

$$B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$$

$$B(v, fw) = fB(v, w)$$

We now define differentiability using the total differential.

Definition 25. A function f is said to be differentiable in $a \in U \subset \mathbb{R}^n$ if $\nabla f(a)$ exists and:

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - \underbrace{\langle \nabla f(a), h \rangle}_{\text{total differential}}}{h} = 0$$

The above definition assures that the total change in f over some variable h occurs quicker than the change in h .

Similarly, we define directional derivatives.

Definition 26. For some $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ the directional derivative at a with respect to $v \in U$ is:

$$D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}$$

Now some non-implications to pay attention to.

Remark 1.3. Non-implications

- If all directional derivatives at a exist, it does not mean that f is differentiable at a .
- Directional derivatives with respect to v are not necessarily equal to $\langle \nabla f, v \rangle$ this is only true when f is differentiable at a .

Definition 27. Notation for higher order derivatives We notate a derivative of order 3 with respect to x, y, x in respective order as:

$$D_{x,y} f(x)$$

Definition 28. Tangent hyperplane equation For some $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ we define the tangent hyperplane at some point $a \in U$ as:

$$f(\bar{x}) = f(a) + \langle \nabla f(a), h \rangle$$

Definition 29. Generalized chain rule Let k be a curve from $\mathbb{R} \rightarrow \mathbb{R}^n$ and f a function from $\mathbb{R}^n \rightarrow \mathbb{R}$ such that:

1. k is differentiable at $t_0 \in \mathbb{R}$
2. f is differentiable at $k(t_0)$

Then we have that:

$$(f \circ k)(t_0)' = \langle \nabla f(k(t_0)), k'(t_0) \rangle$$

Proof. We briefly prove the above theorem here.

First let's note the following: The fact that k is differentiable at t_0 means that

$$k(t_0 + h) - k(t_0) = k'(t_0)h + o(|h|)$$

From this it follows that the LHS is also $O(|h|)$ hence if some $s(h)$ is little o of the LHS, it is definitely little of the RHS. Now we have:

$$\begin{aligned} f(k(t_0 + h) - k(t_0)) - f(k_0) &= \langle \nabla f(k(t_0)), k(t_0 + h) - k(t_0) \rangle + o(|k(t_0 + h) - k(t_0)|) = 0 \\ f(k(t_0 + h) - k(t_0)) - f(k_0) &= h \langle \nabla f(k(t_0)), k'(t_0) \rangle + o(|h|) \end{aligned}$$

□

Now we define higher order derivatives.

Definition 30. A partial derivative of second order is written as:

$$\frac{\partial^2 f}{\partial x_j \partial x_i} \equiv D_{ji}f = \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x_i}(\bar{a} + he_j) - \frac{\partial f}{\partial x_i}(\bar{a})}{h}$$

An interesting result is that the order of partial differentiation does not matter.

Theorem 1.3. If $f \in C^m(U)$ then $\forall r \leq m$ we have that:

$$D_{j_1, \dots, j_r} = D_{i_1, \dots, i_r}$$

where the i are a permutation of the j

Now we introduce the hessian.

Definition 31. *Hessian of a matrix*

$$Hess_f(x) = \begin{bmatrix} D_{11} & \dots & D_{1n} \\ \vdots & \ddots & \vdots \\ D_{n1} & \dots & D_{nn} \end{bmatrix}$$

Using the hessian, we define the second order taylor approximaion of a multivariable functions as below.

Theorem 1.4.

$$f(\bar{x} + \bar{h}) = f(\bar{x}) + \langle \nabla f(\bar{x}), \bar{h} \rangle + \frac{1}{2} \bar{h}^T Hess(\bar{x}) \bar{h}$$

Intuition for the above theorem is easy. Just remember that the $\frac{1}{2}$ comes from the fact that it is a taylor expansion (analogue in 1-d is $f(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(x - a)^2$)

Theorem 1.5. Whenever we have a product, the following holds:

$$\nabla(fg) = \nabla f g + \nabla g f$$

1.7 Vector Fields

Definition 32. *Vector field* A vector field is simply a function from $\mathbb{R}^n \rightarrow \mathbb{R}^m$

How might we imagine the partial derivative of a vector field?

Definition 33. Some $v : U \rightarrow \mathbb{R}^m$ is partially differentiable at $\bar{a} \in U$ if:

$$D_j v_i(a) = \frac{\partial v_i}{\partial x_j}(a) \equiv \lim_{h \rightarrow 0} \frac{v_i(a + he_j) - v_i(a)}{h} \text{ exists}$$

Thus the natural analogue of the Hessian now called the Jacobian of a vector space is (though the Jacobian lists partial derivatives of the first order:

Definition 34. Jacobian

$$J_f(x) = \begin{bmatrix} D_1 v_1 & \dots & D_n v_1 \\ \vdots & \ddots & \vdots \\ D_1 v_n & \dots & D_n v_n \end{bmatrix}$$

Given this definition of the Jacobian, we define differentiability at some point $a \in U \subset \mathbb{R}^n$ as:

Definition 35. Differentiability Some $v : U \rightarrow \mathbb{R}^m$ is differentiable at $a \in U$ if there exists a linear map J_f such that

$$\lim_{h \rightarrow 0} \frac{v(a+h) - v(a) - J_f(h)}{\|h\|} = 0$$

Now the same concept of $C^1(\mathbb{R})$ is applicable to vector fields. Some vector field v is of class $C^1(U)$ if all partial derivatives $D_j v_i : U \rightarrow \mathbb{R}$ exist and are continuous in U .

And now we ask what the Jacobian might be for the composition of two vector spaces?

Theorem 1.6. The composition rule is exactly like the chain rule in single variable calculus. For some $v : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and some $w : \mathbb{R}^m \rightarrow \mathbb{R}^n$ we have

$$J_w \circ v(x) = J_w(v(x))J_v(x)$$

Hence an intuitive way to calculate the i, j position element of $J_w \circ v(x)$ is to calculate:

$$\frac{\partial w_i(v(x))}{\partial x_j} = \sum_{k=1}^n \frac{\partial v_k}{\partial x_j}(x) \frac{\partial w_i}{\partial v_k} v(x)$$

Let's build a better intuition for the Jacobian matrix. It's columns are simply partial derivatives of $v(x)$ with respect to x_i . Our conversely, the rows of the Jacobian are the gradient of i th row of the function. Hence we have with i fixed a constant:

$$J_W(x)_{i,j} = \nabla f(x)_i^T$$

1.8 Implicit function theorem

Consider the real-valued function $f(x, y) = x^2 + e^y - 1$. Now our goal is to determine the level set $N_f(0)$. Hence we write $x^2 + e^y - 1 = 0$ and now our goal is to find y as a function of x . Well as it turns out $y = g(x) = \ln(1 - x^2)$. Now $g'(x) = \frac{-2x}{1-x^2}$ which also happens to be $g'(x) = \frac{-D_1 f(x, g(x))}{D_2 f(x, g(x))}$. Hence for some $U \subset \mathbb{R}^2$ we state the following:

Theorem 1.7. Let $f : U \rightarrow \mathbb{R}$ be of class C^1 (ie. continuous partial derivatives) such that $f(a, b) = 0$ and also $D_2 f(a, b) \neq 0$. Then $\exists B(a, \epsilon)$ and a unique function $g : B(a, \epsilon) \rightarrow \mathbb{R}$ such that:

1. graph of g is in U
2. $g(a) = b$
3. $f(x, g(x)) = 0$
4. $g'(x) = \frac{-D_1 f(x, g(x))}{D_2 f(x, g(x))}$

We may now also generalize to higher dimension:

Theorem 1.8. Implicit function theorem Let $F \in C^1$ such that $f(\bar{x}, y) = 0$ and $\frac{\partial F}{\partial y} \neq 0$. Then there exists a unique $y = f(\bar{x})$ such that

1. graph of f is in U
2. $f(\bar{x}) = y$
3. $f(\bar{x}, y) = 0$
4. $\frac{\partial f}{\partial x_i} = \frac{-\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial y}}$