Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

DEEP LEARNING ON ABSTRACT SIMPLICIAL COMPLEXES

Relatore:
Prof.ssa. Rita Fioresi

Presentata da:
Tommaso Lamma

Anno Accademico 2020/2021

Contents

Chapter 1

Preliminaries on topology

The essential idea in algebraic topology is to convert problems about topological spaces and continuous functions into problems about algebraic objects and their homomorphisms, this way one hopes to end up with an easier problem to solve.

Simplicial Complexes

In this section we shall define structures called simplicial complexes and discuss some of their properties. In order to define these structures we need the definitions of convex hull and affine independence in $\mathbb{R}^n$. In this chapter we recall some notions of algebraic topology, such as simplicial complexes and homology. For more details we invite the reader to consult [3], a good reference also for the preliminary necessary notions of topology we are unable to treat here.

Definition 1.1. Let $A \subset \mathbb{R}^n$, we define A to be *convex* if

$$x, y \in A \Rightarrow tx + (1+t)y \in A$$

for all $t \in [0,1]$.

In Figure 1.1 can see in blue an example of a convex set, since we cannot find to points whose linking segment lies in part ouside the set. Conversely, since the green set cointains two points linked by a segment that partially lies outside the set, we call that set non-convex.

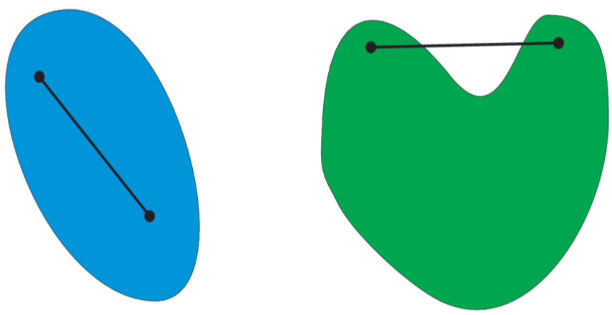

Fig. 1.1: Illustration of a convex (blue) and a non-convex (green) set.

Definition 1.2. Let $\sigma := \{x_i\}_{i\in I}$ be a subset of $\mathbb{R}^n$, where I is a finite set of indexes, we define σ to be *affinely independent* if $\{x_0 - x_i\}_{i\in I\text{-}\{0\}}$ is linearly independent.

We show now that the definition of affine independence of $\sigma = \{x_i\}_{i\in I} \subset \mathbb{R}^n$ is independent of the choice of x_0.

Proposition 1.3. *Let $\sigma := \{x_i\}_{i\in I}$ be a finite subset of $\mathbb{R}^n$, let $j \in I$ then, if $\{x_j - x_i\}_{i\in I\text{-}\{j\}}$ is linearly independent, also $\{x_0 - x_i\}_{i\in I\text{-}\{0\}}$ is.*

Proof. If $j = 0$ the statement is trivially true. Let $j \neq 0$ and $\lambda_i \in \mathbb{R}$ for all $i \neq j$, then

$$\sum_{i\in I\text{-}\{j\}} \lambda_i(x_j - x_i) = 0 \Rightarrow \lambda_i = 0 \quad \forall i \in I\text{-}\{j\}.$$

Let then $\mu_i \in \mathbb{R}$ for all $i \neq 0$, and suppose

$$\sum_{i\in I\text{-}\{0\}} \mu_i(x_0 - x_i) = (x_0 - x_j) \sum_{i\in I\text{-}\{0\}} \mu_i + \sum_{i\in I\text{-}\{0\}} \mu_i(x_j - x_i) = 0.$$

If we define $\mu_0 := -\sum_{i\in I\text{-}\{0\}} \mu_i$ we have that

$$0 = \sum_{i\in I} \mu_i(x_j - x_i) = \sum_{i\in I\text{-}\{j\}} \mu_i(x_j - x_i) \Rightarrow \mu_i = 0 \quad \forall i \in I\text{-}\{j\},$$

which proves our proposition. the definition of affine ind the definition of affine independence is well stated.ependence is well stated. □

Definition 1.4. Let $\sigma := \{x_i\}_{i\in I}$ be a finite subset of $\mathbb{R}^n$, we define the *covex set generated by σ* to be the smallest convex set containing X according to the inclusion relation. We shall denote this set by $[\sigma]$ and call it *convex hull* of σ.

Since the intersection of convex sets is convex, the convex set generated by σ can be equivalently defined as the intersection of all convex sets containing σ.

Theorem 1.5. *Let $\sigma := \{x_i\}_{i\in I}$ be a finite subset of $\mathbb{R}^n$, if σ is affinely independent then the convex set generated by σ is*

$$[\sigma] = \{\sum_{i\in I} \lambda_i x_i : \lambda_i \geq 0, \sum_{i\in I} \lambda_i = 1\}.$$

Furthermore for any point $x \in [\sigma]$ we have that

$$x = \sum_{i\in I} \lambda_i x_i = \sum_{i\in I} \mu_i x_i \Rightarrow \lambda_i = \mu_i \forall i \in I,$$

where $\lambda_i, \mu_i \geq 0$ and $\sum_{i\in I} \lambda_i = \sum_{i\in I} \mu_i = 1$.

Proof. Let $C := \{\bigcap_\alpha C_\alpha : \sigma \subset C_\alpha, C_\alpha$ convex$\}$, we divide the proof in three steps:

(i) $C \subset [\sigma]$.
This is true if $[\sigma]$ is convex and contains σ. The proof that it contains σ is trivial. In fact for every vertex $x_j = \sum_{i\in I} \delta_{ij} x_i$, and $\sum_{i\in I} \delta_{ij} = 1$.
To prove that it is convex we chose two points $a = \sum_{i\in I} a_i x_i, b = \sum_{i\in I} b_i x_i$ where $a_i, b_i \geq 0 \; \forall i \in I$ and $\sum_{i\in I} a_i = \sum_{i\in I} b_i = 1$. For $t \in [0,1]$

$$ta + (1-t)b = t\sum_{i\in I} a_i x_i + (1-t)\sum_{i\in I} b_i x_i = \sum_{i\in I}(ta_i + (1-t)b_i)x_i.$$

Since $ta_i + (1-t)b_i \geq 0$ and $\sum_{i\in I}(ta_i + (1-t)b_i) = t\sum_{i\in I} a_i + (1-t)\sum_{i\in I} b_i = 1$ for all $i \in I$, our statement is proven.

(ii) $[\sigma] \subset C$.
If all but one the λ_i are zero certaintly $\sum_{i\in I} \lambda_i x_i \in C$, since C contains all the vertexes. The inuctive hypothesis, by relabeling, is that if the first $\lambda_0, ..., \lambda_{n-1}$ are non-zero, hence

not even 1, then $\sum_{i\in I}\lambda_i x_i \in C$. We want to show that whenever $\lambda_0,\dots,\lambda_n$ are non-zero then also $\sum_{i\in I}\lambda_i x_i \in C$, since $\lambda_n \neq 1$ we have that

$$\sum_{i\in I}\lambda_i x_i = \sum_{i=0}^{n}\lambda_i x_i = \lambda_n x_n + \sum_{i=0}^{n-1}\lambda_i x_i = \lambda_n x_n + (1-\lambda_n)\sum_{i=0}^{n-1}\frac{\lambda_i}{1-\lambda_n}x_i.$$

Since $\sum_{i=0}^{n-1}\frac{\lambda_i}{1-\lambda_n} = 1$, for the inductive hypothesis $\sum_{i=0}^{n-1}\frac{\lambda_i}{1-\lambda_n}x_i \in C$. Also the vertex x_n is contained in C by definition, therefore, being C convex and $\lambda_n \in [0,1]$, it follows that

$$\lambda_n x_n + (1-\lambda_n)\sum_{i=0}^{n-1}\frac{\lambda_i}{1-\lambda_n}x_i \in C.$$

Accordingly $\sum_{i\in I}\lambda_i x_i \in C$, by induction we conclude the proof.

(iii) $\sum_{i\in I}\lambda_i x_i = \sum_{i\in I}\mu_i x_i \Rightarrow \lambda_i = \mu_i \forall i \in I$.
Let $\sum_{i\in I}\lambda_i x_i = \sum_{i\in I}\mu_i x_i$, then also $x_0\sum_{i\in I}\lambda_i + \sum_{i\in I}\lambda_i(x_i - x_0) = x_0\sum_{i\in I}\mu_i + \sum_{i\in I}\mu_i(x_i - x_0)$, and since both λ_i and μ_i are normalised we have that

$$\sum_{i\in I}(\lambda_i - \mu_i)(x_0 - x_i) = \sum_{i\in I\text{-}\{0\}}(\lambda_i - \mu_i)(x_0 - x_i) = 0 \Rightarrow \lambda_i = \mu_i \quad \forall i \in I\text{-}\{0\},$$

because of the affine independence. □

Definition 1.6. We define a *p-simplex* $[\sigma]$ to be the convex hull of an affinely independent set $\sigma := \{x_i\}_{i\in I} \subset \mathbb{R}^n$, where $p = |I| - 1$ is called dimension of the p-simplex.

Theorem 1.5 gives us the possibility to represent a point in a simplex $[\sigma]$ via a finite set of real parameters defined in the range $[0,1]$ and satisfying the normalisation condition $\sum_{i\in I}\lambda_i = 1$. Such parameters are called *baricentric coordinates* of $[\sigma]$.

The points in σ are called *vertexes* of the simplex $[\sigma]$, accordingly we define the vertex set of a simplex $[\sigma]$ to be $Vert([\sigma]) = \sigma$.

Definition 1.7. Let $[\sigma]$ be a p-simplex and $p,t \in \mathbb{N}$, we say that another t-simplex $[\tau]$ is a *face* of $[\sigma]$ or equivalently that $[\sigma]$ is a *coface* of $[\tau]$, and we write $[\tau] \leq [\sigma]$, if $\tau \subset \sigma$, where $t \leq p$.

Now we are ready for our main definitions.

Definition 1.8. We define a *simplicial complex* $\mathcal{G}$ to be a collection of simplexes such that

(i) if any simplex $[\tau] \leq [\sigma]$ and $[\sigma] \in \mathcal{G}$, then $[\tau] \in \mathcal{G}$,

(ii) if $[\sigma],[\tau] \in \mathcal{G}$, then $[\sigma] \cap [\tau] \in \mathcal{G}$.

1.2 represents a simplicial complex, while 1.3 represents a collection of simplexes which is not a simplicial complex.

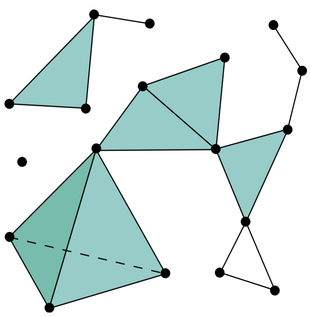

Fig. 1.2: Example of simplicial complex.

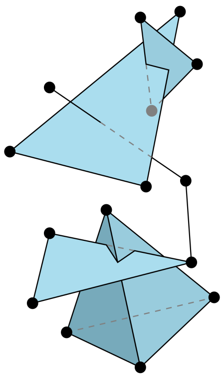

Fig. 1.3: Set of simplexes which is not a simplicial complex.

In fact, in Figure 1.3 that the intersection property of simplicial complexes is not satisfied.

In this first part we analyzed simplicial complexes as peculiar subsets of $\mathbb{R}^n$, which from now on we shall call *geometric simplicial complexes*. Although this approach provides simplicial complexes with the topology inherited from the metric space, it hides the power of simplicial complexes to describe networks and interactions which exist independently of that topology. To make this distinction clear we will treat simplicial complexes as a realization of more abstract objects called *abstract simplicial complexes*. A richer discussion of abstract simplicial complexes can be found in [4, 2].

Definition 1.9. Let $\mathscr{V}$ be a finite set, we define an *abstract simplicial complex* $\mathscr{A}$ to be a family of non empy subsets of $\mathscr{V}$ such that:

(i) if $v \in \mathscr{V}$, then $\{v\} \in \mathscr{A}$,

(ii) if $\sigma \in \mathscr{A}$ and $\tau \subset \sigma$, then $\tau \in \mathscr{A}$.

We call the member of this family *abstract simplexes*.

One calls $\mathscr{V}$ the *vertex set* of $\mathscr{A}$ and denotes it by $Vert(\mathscr{A})$; since the vertex set is finite we expect every abstract simplex to be also finite, therefore we might use the notation $\sigma = \{v_i\}_{i\in I_\sigma}$, which so far we consider invariant under arbitrary permutations of the finite index set I_σ.

Definition 1.10. Let $\mathscr{A}$ be an abstract simplicial complex and $\mathscr{G}$ a geometric simplicial complex, if for all $\{x_i\}_{i\in I} \in \mathscr{A}$ also $[x_i]_{i\in I} \in \mathscr{G}$ we say that $\mathscr{G}$ is a *geometric realization* of $\mathscr{A}$.

While every geometric simplicial complex can be thought as a geometric realization of an abstract simplicial complex, the existence of a geometric realization for an arbitrary abstract simplicial complex is not trivial at all.

Theorem 1.11. *Let $\mathscr{A}$ be an n-dimensional abstract simplicial complex, then it admits a geometric realization in* $\mathbb{R}^{2n+1}$.

A proof of this theorem can be found in [2].

Both for abstract and geometric simplicial complexes one can define maps called *simplicial maps* in order to obtain a category whose equivalences are called isomorphisms. The geometric realisation is unluckily not a functor on the whole category of abstract simplicial complexes due to the limitation imposed by 1.11. Nevertheless, when they exist, the geometric realizations of isomorphic abstract simplicial complexes are theirselves isomorphic. A short discussion of category theory can be found in Appendix A.

In the following sections we shall use abstract simplicial complexes, which can be always thought geometrically in the appropriate euclidean space.

Conclusion

Most of the deep learning techniques used today are based on models which learn a partition of the set of smooth functions defined on euclidean domains into human friendly equivalence classes. Although this approach has been successful in modern machine learning, it only deals with a really small set of domains. The goal of geometric deep learning is to extend this method to data defined on manifolds and simplicial complexes.
Convolution on euclidean domains is itself based on the translation invariance of such domains. In fact the convolution of a function $f:\mathbb{R}^n \to \mathbb{R}$, with some filter $g:\mathbb{R}^n \to \mathbb{R}$ is

$$(f * g)(x) = \langle f, g \circ T^{-1} \rangle_{L^2},$$

where $T:\mathbb{R}^n \to \mathbb{R}^n$ is a translation represented by the vector x. Moreover, one could also consider such a convolution to be defined on the translation group itself, represented by some $\mathbb{R}^n$. In the same way one can define a convolution on the group $(\mathbb{Z},+)$ as $(a*b)_n = \sum_{i\in\mathbb{Z}} a_i b_{n-i}$. Similarly an intersting example is that of images, which are samples ad grids, any image can be thought as a function defined on the group $(\mathbb{Z}\times\mathbb{Z},+)$. Such definitions of convolution operators are equivariant with respect to the action of the group they are defined upon. In image recognition translation equivariance is necessary, nevertheless the most common CNN's need to learn the rotations of the same filter as different filters in order to become rotation equivariant. Although manifolds and simplicial complexes are not in general groups, G-equivariant CNN's (see [1]) could be the key to reveal the secrets behind the success of such architectures.

Appendix A

Category Theory

At the beginning of chapter 1 we mentioned that a relevant concept in algebriac topology is that of conversion form structures where a problem cannot be solved to structures where the problem can be solved. Not all conversions are good ones, we need the language of category theory to define the rules that make a conversion a good conversion, i.e. functoriality.

Definition A.1. A *category* $\mathbf{C}$ consists of three ingerdients:
a class of *objects* $Obj(\mathbf{C})$; sets of *morphisms* $Hom(A,B)$ for every ordered pair $(A,B) \in Obj(\mathbf{C}) \times Obj(\mathbf{C})$; a composition $Hom(A,B) \times Hom(B,C) \to Hom(A,C)$, denoted by $(f,g) \mapsto f \circ g$ for every $A,B,C \in Obj(\mathbf{C})$, satisfying the following axioms:

(i) the family of $Hom(A,B)$ is pairwise disjoint,

(ii) the composition, when defined, is associative,

(iii) for each $A \in Obj(\mathbf{C})$ there exists an *identity* $1_A \in Hom(A,A)$ such that for $f \in Hom(A,B)$ and $g \in Hom(C,A)$ we have that $1_A \circ f = f$ and $g \circ 1_A = g$.

Instead of writing $f \in Hom(A,B)$, we usually write $f : A \to B$.

Definition A.2. Let $\mathbf{A}$ and $\mathbf{C}$ be categories, a *functor* $T : \mathbf{A} \to \mathbf{C}$ is a function, that is,

(i) $A \in Obj(\mathbf{A}) \implies TA \in Obj(\mathbf{C})$,

(ii) $f : A \to A' \implies Tf : TA \to TA' \quad A,A' \in Obj(\mathbf{A})$,

(iii) if f,g are morphisms in $\mathbf{A}$ for which $g \circ f$ is defined, then $T(g \circ f) = (Tg) \circ (Tf)$,

(iv) $T(1_A) = 1_{TA} \quad \forall A \in \mathbf{A}$.

The property (iii) of the previous definition actually defines what we shall call *covariant functors*. If instead we require $T(g \circ f) = (Tf) \circ (Tg)$, we are defining a so called *contravariant functor*.

Definition A.3. An *equivalence* in a category $\mathbf{C}$ is a morphism $f : A \to B$ for which there exists a morphism $g : B \to A$ such that $f \circ g = 1_B$ and $g \circ f = 1_A$.

Theorem A.4. *If $\mathbf{A}$ and $\mathbf{C}$ are categories and $T : \mathbf{A} \to \mathbf{C}$ is a functor of either variance, then whenever f is an equivalence on $\mathbf{A}$ then Tf is an equivalence on $\mathbf{C}$.*

Proof. We apply T to the equations $f \circ g = 1_B$ and $g \circ f = 1_A$, that for a covariant functor leads to $(Tf) \circ (Tg) = T(1_B) = 1_{TB}$ and $(Tg) \circ (Tf) = T(1_A) = 1_{TA}$. □

A category that will be used in the following section is the category of topological spaces and continuous functions.

Proposition A.5. *Topological spaces and continuous functions are a category **Top**, whose equivalences are called homeomorphisms.*

Other examples of categories can be found in [4].

Bibliography

[1] Taco S. Cohen and Max Welling. *Group Equivariant Convolutional Networks*. 2016.

[2] Herbert Edelsbrunner and John Harer. *Computational Topology - an Introduction.* American Mathematical Society, 2010.

[3] Allen Hatcher. *Algebraic Topology*. Cambridge Univ. Press, 2000.

[4] Joseph J. Rotman. *An Introduction to Algebraic Topology*. Graduate Texts in Mathematics. Springer-Verlag New York, 1988.

[5] I. M. Singer and J. A. Thorpe. *Lecture Notes on Elementary Topology and Geometry*. Springer, New York, NY, 1976.