ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

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

FRONTIERS OF DEEP LEARNING

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.

1.1 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 [4], a good reference also for the preliminary necessary notions of topology we are unable to treat here.

Definition 1.1.1. 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: every segment joining two points of the set lies within the set. The green set is not convex, in fact we see that the segment in the illustration partially lies outside of the set.

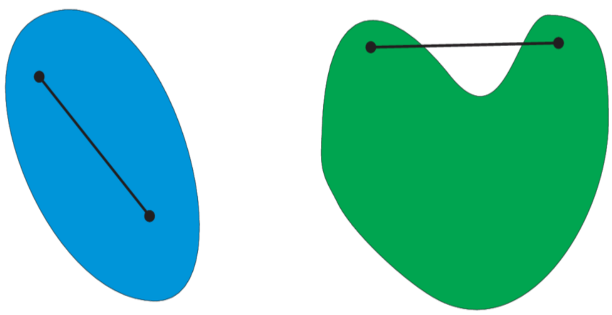

Fig. 1.1: Illustration of a convex (blue) and a non-convex (green) set.

Definition 1.1.2. 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 1.1.3. *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. □

Definition 1.1.4. Let $\sigma := \{x_i\}_{i\in I}$ be a finite subset of $\mathbb{R}^n$, we define the *convex 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 1.1.5. *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 1.1.6. 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 1.1.5 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 1.1.7. 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 $[\tau] \le [\sigma]$, if $\tau \subset \sigma$, where $t \le p$.

Now we are ready for our main definitions.

Definition 1.1.8. We define a *simplicial complex* $\mathcal{G}$ to be a collection of simplexes such that

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

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

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

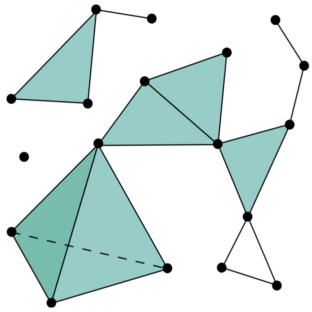

Fig. 1.2: Example of simplicial complex.

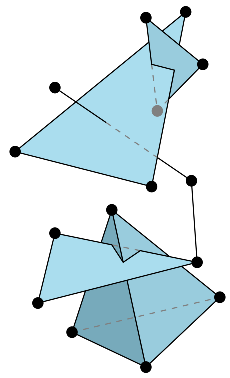

Fig. 1.3: Set of simplexes which is not a simplicial complex.

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

1.2 Abstract Simplicial Complexes

In section 1.1 we studied simplicial complexes as subsets of $\mathbb{R}^n$. From now we shall call them *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 [3] at 3.1 or in [8] at 7.3.

Definition 1.2.1. 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 members of this family *abstract simplexes*.

We call $\mathscr{V}$ the *vertex set* of $\mathscr{A}$ and denote it by $Vert(\mathscr{A})$; since the vertex set is finite, every abstract simplex is finite, therefore we can use the notation $\sigma = \{v_i\}_{i\in I}$, to denote a simplex in $\mathscr{A}$.

Definition 1.2.2. 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 1.2.3. *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 [3] at 3.1.

Both for abstract and geometric simplicial complexes one can define maps called *simplicial maps*. We obtain a category whose equivalences are called isomorphisms. 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 $\mathbb{R}^{2d+1}$.

1.3 Simplicial Homology

An important field in algebraic topology is homology theory. We shall discuss homology theory to the extent that allows us to define the laplacian operator on simplicial complexes, for supplementary readings see [9] at 6.1 or [4] at 2.1. First we want to equip our simplicial complexes with an orientation. So far we have considered the simplex $\{x_i\}_{i\in I}$ to remain unchanged under reorderings of the index set I, but in most applications this is not the case.

Proposition 1.3.1. *Let $\{x_i\}_{i\in I}$ be a p-simplex and*

$$\{x_i\}_{i\in I} \sim \{x_i\}_{i\in\pi(I)} \iff sgn(\pi) = 1,$$

where $\pi : I \to I$ is a permutation of the indexes and $sgn(\pi)$ its sign, then $\sim$ is an equivalence relation.

Remark 1.3.2. Let S_{p+1} be the group of permutations of a p-simplex, and {-1,1} a multiplicative group, we recall the fact that $sgn : S_{p+1} \to \{\text{-}1,1\}$ is a group homomorphism, that is

$$sgn(\pi\eta) = sgn(\pi)sgn(\eta) \quad \forall \pi,\eta \in S_{p+1}.$$

Proof. We divide the proof in three steps:

(i) Since $sgn(id_I) = 1$ we have that $\{x_i\}_{i\in I} \sim \{x_i\}_{i\in id_I(I)} = \{x_i\}_{i\in I}$.

(ii) Since $sgn(\pi)sgn(\pi^{-1}) = sgn(\pi\pi^{-1}) = sgn(id_I) = 1$ we have that $sgn(\pi) = sgn(\pi^{-1})$, therefore $\{x_i\}_{i\in I} \sim \{x_i\}_{i\in\pi(I)} \iff \{x_i\}_{i\in\pi(I)} \sim \{x_i\}_{i\in I}$.

(iii) The transitivity is a consequence of the fact that the product of two even permutations is also even. □

Definition 1.3.3. We define an *oriented simplex* $|x_i\rangle_{i\in I}$ to be a simplex $\{x_i\}_{i\in I}$ together with the choice of one of the two equivalence classs with respect to $\sim$.

Mathematical Foundation of CNN's

Definition 1.3.4. Let $\mathscr{A}$ be a simplicial complex, we define the *group of p-chains*

$$C_p(\mathscr{A},\mathbb{R}) := \frac{\{\sum_{i\in I}\lambda_i|\sigma_i\rangle : \lambda_i \in \mathbb{R}\}}{\{|x_0,x_1,\dots,x_p\rangle + |x_1,x_0,\dots x_p\rangle\}}.$$

In the previous definition the sum $\sum_{i\in I}\lambda_i|\sigma_i\rangle$ is a formal sum that can be generalized from $\mathbb{R}$ to an arbitrary abelian group G, for a deeper insight into free abelian groups see [6] at 1.7. Nevertheless, for the most of the applications the groups $\mathbb{Z},\mathbb{R},\mathbb{Z}_2$ are considered. To keep our notation light we shall write C_p instead of $C_p(\mathscr{A},\mathbb{R})$.

Proposition 1.3.5. *$C_p(\mathscr{A},\mathbb{R})$ is an abelian group. In particular it is a vector space on $\mathbb{R}$.*

A particularly relevant role in homology theory is played by the *boundary map*. First we define the boundary of an oriented simplex.

Definition 1.3.6. Let $|\sigma\rangle = |x_0, ..., x_{p+1}\rangle$ be an oriented $(p+1)$-simplex. The boundary $\partial|\sigma\rangle$ of $|\sigma\rangle$ is the p-chain defined by

$$\partial_{p+1}|\sigma\rangle := \sum_{i=0}^{p+1}(-1)^i|x_0, ..., \hat{x_i}, ..., x_{p+1}\rangle$$

where the ˆ over a symbol means that symbol is deleted.

Remark 1.3.7. Note that whenever we are able to construct a geometric realization for the oriented simplicial complex, the set $\bigcup_{i=0}^{p+1}[x_0, ..., \hat{x_i}, ..., x_p + 1]$ is the topological boundary of $[\sigma]$.

Furthermore we are able to extend the boundary from simplexes to chains.

Definition 1.3.8. We define the *boundary map* $\partial_{p+1} : C_{p+1} \to C_p$ to be the group homomorphism defined by

$$\partial_{p+1}(\sum_{i\in I}\lambda_i|\sigma_i\rangle) := \sum_{i\in I}\lambda_i\partial_{p+1}|\sigma_i\rangle.$$

An important property of boundary maps is the following, which we shall denote as *homology lemma*.

Lemma 1.3.9. *The boundary maps satisfy* $\partial_p \circ \partial_{p+1} = 0$.

Proof. Since the boundary maps are linear it is sufficient to check this on the generators. Let $\partial_{p+1}|x_0, \ldots, x_{p+1}\rangle = \sum_{i=0}^{p+1}(-1)^i|x_0, \ldots, \hat{x_i}, \ldots, x_{p+1}\rangle$ then

$$(\partial_p \circ \partial_{p+1})|x_0, \ldots, x_{p+1}\rangle = \sum_{j=0, j\neq i}^{p+1}\sum_{i=0}^{p+1}(-1)^{i+j}|x_0, \ldots, \hat{x_i}, \ldots, \hat{x_j}, \ldots, x_{p+1}\rangle = 0. \qquad \square$$

The homology lemma is necessary to define the *homology group*.

Definition 1.3.10. We define the *p-homology group* to be

$$H_p := \frac{ker\partial_p}{im\partial_{p+1}},$$

where $im\partial_{p+1}$ is called the group of simplicial *p-cycles* and $ker\partial_p$ is called the group of simplicial *p-boundaries*.

The homology group is therefore, intuitively, the space of cycles that are not boundaries. Without the homology lemma the quotient would not well defined since $im\partial_{p+1} \subset ker\partial_p$ would not be satisfied.

1.4 Simplicial Cohomology

In order to define data on simplicial complexes we are interested in studying the dual of the chains $|\sigma\rangle$, which we shall call cochains $\langle\sigma|$.

Definition 1.4.1. Let C_p be the group p-chains, we define the group of *p-cochains* to be $C^p := Hom(C_p, G)$, i.e. the dual space of C_p.

As well as with for chains, we have also for cochains a sequence of homeomorphisms called *coboundary maps*. The coboundary maps are defined to be the dual of the boundary maps, hence satisfying a dual version of the homology lemma called *cohomology lemma*.

Definition 1.4.2. The dual of the boundary maps which we shall call *coboundary maps*, is the group homomorphsm defined by

$$d_{p+1} : C^p \to C^{p+1} \quad d_p \langle\sigma| := \langle\sigma|\partial_{p+1} \quad \forall\langle\sigma| \in C^p.$$

Therefore $(d_{p+1}\langle\sigma|)|\tau\rangle = \langle\sigma|\partial_{p+1}|\tau\rangle \quad \forall|\tau\rangle \in C_p$.

The proof of $d^p + 1 \circ d_p$ follow directly from the homology lemma. The cohomology lemma allows us to define the *cohomology group*.

Definition 1.4.3. We define the *p-cohomology group* to be

$$H^p := \frac{ker d_{p+1}}{im d_p},$$

where $im d_{p-1}$ is the group of simplicial p-cocycles and $ker d_p$ is the group of simplicial p-coboundaries.

To have a specular intuition of what the coboundary operator actually does, we equipe our simplicial chain groups with an inner product. This finite dimensional Hilbert space structure shall allow us to represent the cochain $\langle\sigma|$ evaluated on the chain $|\tau\rangle$ as the inner product $\langle\sigma|\tau\rangle$.

Definition 1.4.4. We define the scalar product $\langle|\rangle : C_p \times C_p \to \mathbb{R}$ on the canonical basis of C_p to be

$$\langle i|j\rangle = \delta_{ij},$$

where i, j are any two p-simplexes in the canonical basis of C_p and δ_{ij} is the Kronecker Delta.

The same Hilbert space structure allows us to represent the coboundary operator d acting on a cochain with its cochain representative.

Definition 1.4.5. Let $d_{p+1} : C^p \to C^{p+1}$ be the coboundary operator, then for any $\langle\sigma| \in C^p, |\tau\rangle \in C_{p+1}$ we can define its dual representation $\partial^\dagger_{p+1}$ by

$$(d_{p+1}\langle\sigma|)|\tau\rangle = \langle\sigma|\partial_{p+1}|\tau\rangle = \langle\tau|\partial^\dagger_{p+1}|\sigma\rangle.$$

We can notice that our definitions lead to the restatement of the equivalent of the generalized Stokes' theorem on simplicial complexes according to the integration previously defined, i.e. $(d\langle\sigma|)|\tau\rangle = \langle\sigma|\partial|\tau\rangle$. One can therefore think of the coboundary operator as a discrete exterior derivative acting on cochains. Similarly we name the inner product *integration* due to the analogy previously stated, the choice of a measure of integration is equivalent to a choice of a basis for the chain group.

Chapter 2

Graphs

2.1 Homology and cohomology on graphs

In chapter 1 we studied homology and cohomology on abstract simplicial complexes, in this chapter we shall discuss in detail 1-dimensional abstract simplicial complexes.

Definition 2.1.1. A *graph* is a 1-dimensional abstract simplicial complex.

In a graph $\mathscr{G}$ the non trivial chain groups are therefore $C_0(\mathscr{G})$ and $C_1(\mathscr{G})$, also called vertex and edge spaces. The canonical basis of the vertex space is the set of vertexes $\{|i\rangle\}_{i\in I}$, and that of edges which is a subset of $\{|i,j\rangle\}_{(i,j)\in I\times I}$

Definition 2.1.2. Let $\mathscr{G}$ be a graph, we define the *adjacency matrix* A of $\mathscr{G}$ on the canical vertex basis by

$$a_{ij} = \begin{cases} 1 & |i,j\rangle \in C_1 \\ 0 & otherwise \end{cases}$$

Only one boundary ∂_1 and one coboundary $\partial_1^\dagger$ can be defined on graphs, we shall name their dual representatives gradient $d_1^\dagger$ and divergence d_1 respectively.

Definition 2.1.3. Let $\mathscr{G}$ be a graph, we define the *boundary* $\partial_1 : C_1 \to C_0$ by

$$\partial_1|i,j\rangle := |i\rangle - |j\rangle \quad \forall |i,j\rangle \in C_1.$$

Definition 2.1.4. Let $\mathscr{G}$ be a graph, we define the *coboundary* $\partial_1^\dagger : C_0 \to C_1$ by

$$\partial_1^\dagger|i\rangle := \sum_{|i,j\rangle\in C_1} |i,j\rangle.$$

Disegno che mostra il cobordo di un nodo

The prevously defined generalizd Stokes' theorem is on graphs a discrete version of the divergence theorem.

Theorem 2.1.5. *Let $\mathscr{G}$ be a graph, let $|\phi\rangle \in C_0, |\psi\rangle \in C_1$ then $\langle\phi|\partial_1|\psi\rangle = \langle\psi|\partial_1^\dagger|\phi\rangle = (d_1\langle\phi|)|\psi\rangle$.*

Therefore the integral of the cochain $|\psi\rangle$ over the boundary of $|\phi\rangle$ equals the integral of the divergence of $|\psi\rangle$ over $|\phi\rangle$.

Disegno che mostra il teorema della divergenza discreto

On a graph that admits non trivial gradient and divergence, we can define a non trivial laplacian operator as the gradient of the divergence. Although this might sound strange, the inversion of gradient and divergence in the definition of laplacian is due to the fact that we are representing the functions on the graph (cochains) by their dual chains.

Definition 2.1.6. Let $\mathscr{G}$ be a graph, we define the 1-*laplacian* $\Delta_1 : C_0 \to C_0$ by

$$\Delta_1 := \partial_1 \circ \partial_1^\dagger.$$

One intersting propertiy of the 1-laplacian is that the dimension of its kernel equals the connected components of the graph.
Let $\mathscr{G}$ be the graph in Fig..., then the laplacian expressed in terms of the canoncal vertex basis $\{|i\rangle\}_{i\in I}$, is

$$\begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 3 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

The laplacian matrix is calculated as follows

$$\Delta_1|1\rangle = \partial_1|1,2\rangle + \partial_1|1,3\rangle = 2|1\rangle - |2\rangle - |3\rangle,$$

$$\Delta_1|2\rangle = \partial_1|2,1\rangle + \partial_1|2,3\rangle = -|1\rangle + 2|2\rangle - |3\rangle,$$

$$\Delta_1|3\rangle = \partial_1|3,1\rangle + \partial_1|3,2\rangle = -|1\rangle + -|2\rangle + 2|3\rangle,$$

$$\Delta_1|4\rangle = \partial_1|4,5\rangle = |4\rangle - |5\rangle,$$

$$\Delta_1|5\rangle = \partial_1|5,4\rangle + \partial_1|5,6\rangle + \partial_1|5,7\rangle = -|4\rangle + 3|5\rangle - |6\rangle - |7\rangle,$$

$$\Delta_1|6\rangle = \partial_1|6,5\rangle = -|5\rangle + |6\rangle,$$

$$\Delta_1|7\rangle = \partial_1|7,5\rangle = -|5\rangle + |7\rangle,$$

$$\Delta_1|8\rangle = \partial_1|8,9\rangle = |8\rangle - |9\rangle,$$

$$\Delta_1|9\rangle = \partial_1|9,8\rangle = -|8\rangle + |9\rangle.$$

Three invariant subspaces emerge from the laplacian, that determine three different laplacians, namely

$$\Delta_1 = \Delta_1^{\mathscr{A}} \oplus \Delta_1^{\mathscr{B}} \oplus \Delta_1^{\mathscr{C}}.$$

Furthermore any of those three blocks has a 1-dimensional kernel, in fact the dimensional equiations for the laplacians are

$$dim(ker\Delta_1^{\mathscr{A}}) = dimC_0(\mathscr{A}) - rank\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = 3 - rank\begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 2 = 1,$$

$$dim(ker\Delta_1^{\mathscr{B}}) = dimC_0(\mathscr{B}) - rank\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} = 4 - rank\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 4 - 3 = 1,$$

$$dim(ker\Delta_1^{\mathscr{C}}) = dimC_0(\mathscr{C}) - rank\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = 2 - rank\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} = 2 - 1 = 1.$$

We can also define an higher dimensional laplacian on the graph.

Definition 2.1.7. Let $\mathcal{G}$ be a graph, we define the 2-*laplacian* $\Delta_1 : C_0 \to C_0$ by

$$\Delta_1 := \partial_1^\dagger \circ \partial_1.$$

One intersting propertiy of the 2-laplacian is that the dimension of its kernel equals the number of independent cycles. In fact we can expand $\Delta_2^{\mathscr{A}} := \partial_1^\dagger \partial_1$ the basis $\{|1,2\rangle, |2,3\rangle, |3,1\rangle\}$ as

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

The laplacian matrix is calculated as follows

$$\Delta_2^{\mathscr{A}} |1,2\rangle = \partial_1^\dagger (|1\rangle - |2\rangle) = 2|1,2\rangle - |3,1\rangle - |2,3\rangle,$$

$$\Delta_2^{\mathscr{A}} |2,3\rangle = \partial_1^\dagger (|2\rangle - |3\rangle) = -|1,2\rangle + 2|2,3\rangle - |3,1\rangle,$$

$$\Delta_2^{\mathscr{A}} |3,1\rangle = \partial_1^\dagger (|3\rangle - |1\rangle) = -|1,2\rangle + 2|3,1\rangle - |2,3\rangle.$$

We can notice that Δ_2 has a 1-dimensional kernel.

$$dim(ker\Delta_2^{\mathscr{A}}) = dimC_1(\mathscr{A}) - rank \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = 3 - rank \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 2 = 1.$$

Since the sum of the three raws is 0 we can say that the only linearly independent 1-cycle is $|1,2\rangle + |2,3\rangle + |3,1\rangle$.
This theorem will be proven in chapter 3.

Chapter 3

Geometric Deep Learning

3.1 Group Equivariant Neural Networks

In many classification task the neural networks seem to learn some symmetries under which the classification is invariant. In some simple situations, where these symmetries are obvious and understandable by humans, we hope to find a neural architecture that already knows the symmetry of the data before starting to learn. An approach was proposed in [**G-CNN**] with group equivariant convolutional networks. Before discussing those particular networks we analyse the meaning of group equivariance.

3.2 Convolutional Neural Networks

In section 3.1 we discussed group equivariant convolutional networks, now we analyse standard convolutional networks as shift equivariant networks.

3.3 Simplicial Neural Networks

An important role in the definition of a convolution on simplicial complexes is played by the Laplacian operator, especially by its eigenfunctions and spectrum. A more thorough discussion of this operator can be found in [5].

Definition 3.3.1. We define the *p-Laplacian* operator to be

$$\Delta_p := \partial_{p+1}\partial^\dagger_{p+1} + \partial^\dagger_p \partial_p =: \Delta^+_p + \Delta^-_p.$$

The Laplacian operator is defined to be self-adjoint and positive definite.

Proposition 3.3.2. *Let Δ_p be a laplacian operator, then $\Delta^\dagger_p = \Delta_p$.*

Proof. Let $|\sigma\rangle, |\tau\rangle \in C_p$

$$\begin{aligned}\langle\sigma|\Delta_p|\tau\rangle &= \langle\sigma|(\partial_{p+1}\partial^\dagger_{p+1} + \partial^\dagger_p\partial_p)|\tau\rangle = \\ &= \langle\tau|(\partial_{p+1}\partial^\dagger_{p+1} + \partial^\dagger_p\partial_p)^\dagger|\sigma\rangle = \\ &= \langle\tau|(\partial_{p+1}\partial^\dagger_{p+1} + \partial^\dagger_p\partial_p)|\sigma\rangle = \langle\tau|\Delta_p|\sigma\rangle.\end{aligned}$$

□

According to the spectral theorem there exists a basis of eigenchains of the Laplacian, and since all $\Delta_p, \Delta^+_p, \Delta^-_p$ are self-adjoint we can say that they all admit a basis of eigenchains.

Proposition 3.3.3. *Let $\Delta_p|\sigma\rangle = \lambda_\sigma|\sigma\rangle$ then $\lambda_\sigma \geq 0$.*

Proof. Let $\Delta_p^+|\sigma\rangle = \lambda_\sigma^+|\sigma\rangle$, we see that $\langle\sigma|\Delta_p^+|\sigma\rangle = |\partial_{p+1}|\sigma\rangle|^2 \geq 0$, and since $\langle\sigma|\Delta_p^+|\sigma\rangle = \lambda_\sigma^+\langle\sigma|\sigma\rangle$ we have that $\lambda+_\sigma \geq 0$.
Let then $\Delta_p^-|\sigma\rangle = \lambda_\sigma^-|\sigma\rangle$, we see that $\langle\sigma|\Delta_p^-|\sigma\rangle = |\partial_p|\sigma\rangle|^2 \geq 0$, and since $\langle\sigma|\Delta_p^-|\sigma\rangle = \lambda_\sigma^-\langle\sigma|\sigma\rangle$, we also have that $\lambda-_\sigma \geq 0$.
Furthermore, since $\Delta_p^+\Delta_p^- = \Delta_p^-\Delta_p^+ = 0$ we have that $[\Delta_p^+, \Delta_p^-] = 0$, thence $[\Delta_p, \Delta_p^\pm] = 0$, therefore $\Delta_p, \Delta_p^+, \Delta_p^-$ share a basis of eigenchains. Let $|\sigma\rangle$ be in that common basis then $\Delta_p|\sigma\rangle = \lambda_\sigma|\sigma\rangle$, where $\lambda_\sigma = \lambda_\sigma^+ + \lambda_\sigma^- \geq 0$. □

Another really interesting property that was first proven by Beno Eckmann in 1944, is that the kernel of the p-Laplacian is isomorphic to the p-homology group.

Theorem 3.3.4. *Let Δ_p be a laplacian operator, then $ker\Delta_p \simeq H_p$.*

Proof. We recall the definition of laplacian $\Delta_p =: \partial_{p+1}\partial_{p+1}^\dagger + \partial_p^\dagger\partial_p$.
Because of the homology lemma $\Delta_p^+\Delta-_p = \Delta_p^-\Delta_p^+ = 0$, therefore $ker\Delta_p^\pm \subset im\Delta_p^\mp$.
It is trivial to see that, since $ker\Delta_p = ker\Delta_p^+ \cap ker\Delta_p^-$, we have that $ker\partial_{p+1}^\dagger \cap ker\partial_p \subset ker\Delta_p$.
Less trivial is the opposite inclusion, in fact, let $|\sigma\rangle \in ker\Delta_p$, we have that $\partial_{p+1}^\dagger|\sigma\rangle \in ker\partial_{p+1} = (im\partial_{p+1}^\dagger)^\perp$ and of course that $\partial_{p+1}^\dagger|\sigma\rangle \in im\partial_{p+1}^\dagger$.
Therefore, since our integration is an inner product, because of the adjunction relation we have that the chain $\partial_{p+1}^\dagger|\sigma\rangle$ must be the null chain, hence the inclusion $\ker\Delta_p \subset \ker\partial_{p+1}^\dagger$. Similarly we can see that $\ker\Delta_p \subset \ker\partial_p$, therefore $ker\Delta_p \subset ker\partial_{p+1}^\dagger \cap ker\partial_p$. Now, since $ker\Delta_p = ker\partial_{p+1}^\dagger \cap ker\partial_p$, we conclude that $ker\Delta_p = (im\partial_{p+1}^\dagger)^\perp \cap ker\partial_p \simeq H_p$. □

Let $\{|e_i\rangle\}_{i\in I}$ be the canonical basis of chains for C_p, namely the chains corresponding to the p-simplexes of the simplicial complex. Nevertheless can also choose the laplacian eigenchains $\{|i\rangle\}_{i\in I}$, such that $\Delta_p|i\rangle = \lambda_i|i\rangle$, as a basis for C_p. This change of basis is reflected on a change of coordinates called *simplicial Fourier transform*.

Definition 3.3.5. Let $|f\rangle \in C_p$, where $dimC_p = n_p$, we define the *simplicial Fourier transform* $\mathscr{F}_p : \mathbb{R}^{n_p} \to \mathbb{R}^{n_p}$ to be

$$(\langle e_i|f\rangle)_{i\in I} \mapsto (\langle i|f\rangle)_{i\in I},$$

where $I = \{1, ..., n_p\}$.

This transform only defines a change of basis since $|f\rangle = \sum_{i\in I}\langle e_i|f\rangle|e_i\rangle = \sum_{i\in I}\langle i|f\rangle|i\rangle$, and therefore is invertible. We can in fact represent the simplicial Fourier transform with the matrix $F_{ij}^{-1} = F_{ij}^\dagger := \langle i|e_j\rangle$, and its inverse $F_{ij} := \langle e_i|j\rangle$. To define a convolution between two p-chains we use the famous convolution theorem $\mathscr{F}(f * \psi) = \mathscr{F}(f)\mathscr{F}(\psi)$. The laplacian is diagonalized by the simplicial Fourier transform, i.e. $\Delta_p = F diag(\lambda_i)_{i\in I}F^\dagger$.

Definition 3.3.6. Let $\{|i\rangle\}_{i\in I}$ be a basis such that $\Delta_p|i\rangle = \lambda_i|i\rangle$, let $|f\rangle, |\psi\rangle \in C_p$, we define the representatives of $|f * \psi\rangle$ on the laplacian eigenchains to be

$$\langle i|f * \psi\rangle := \langle i|f\rangle\langle i|\psi\rangle \quad \forall i \in I.$$

Therefore $|f * \psi\rangle = \sum_{i\in I}\langle i|f\rangle\langle i|\psi\rangle|i\rangle$.

The filters used in [2] are low degree polynomials in the frequency domain, for instance a filter defined by $\langle i|\psi_\mu\rangle = \sum_{n=0}^N \mu_n\lambda_i^n$. This way one can easily define the convolution on the canonical basis as

$$\langle e_i|f * \psi_\mu\rangle = \sum_{j\in I}\langle e_i|j\rangle\langle j|f * \psi_\mu\rangle = \sum_{n=0}^N \mu_n\langle e_i|(\sum_{j\in I}|j\rangle\lambda_j^n\langle j|)|f\rangle = \sum_{n=0}^N \mu_n\langle e_i|\Delta_p^n|f\rangle,$$

therefore $|f * \psi_\mu\rangle = \sum_{n=0}^N \mu_n\Delta_p^n|f\rangle$.

3.4 Message Passing Neural Networks

In the previous section we saw one possible definition of convolutional layers on graph and simplicial complexes [1], here we shall discuss on graphs a more general definition based on the diffusion equation on graphs.

[1] In general on any chain complex

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

Definition A.1. A *category* $\mathbf{C}$ consists of three ingerdients:

1. a class of *objects* $Obj(\mathbf{C})$,
2. sets of *morphisms* $Hom(A,B)$ for every ordered pair $(A,B) \in Obj(\mathbf{C}) \times Obj(\mathbf{C})$,
3. 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) for each $A \in Obj(\mathbf{A})$ it assigns $TA \in Obj(\mathbf{C})$,

(ii) for each morphism $f : A \to A'$ it assigns a morphism $Tf : TA \to TA' \quad \forall A, A' \in Obj(\mathbf{A})$,

(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 *contravariant functor*.

Definition A.3. An *equivalence* in a category $\mathbf{C}$ is a morphism $f : A \to B$ for all $A,B \in Obj(\mathbf{C})$ 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 [8] at 0.3 and in [7].

Bibliography

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

[2] S. Ebli, M. Defferrard, and G. Spreemann. *Simplicial Neural Networks*. 2020.

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

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

[5] Danijela Horak and Juergen Jost. *Spectra of combinatorial Laplace operators on simplicial complexes*. 2011.

[6] Serge Lang. *Algebra*. Springer, New York, NY, 1955.

[7] Saunders MacLane. *Categories for the Working Mathematician*. 1971.

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

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