Canonical form

Notation. *We will consider all complexes to be over a field $\mathbb{K}$ that will either be $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$*

1.1 Ordered complex and canonical form

Definition 1.1 ($\mathbb{R}$-Filtered complex)**.** Let $\{C_k\}_{k=0}^{\infty}$ be a complex. A $\mathbb{R}$-filtration (sometimes we will simply say a filtration), on $\{C_k\}_{k=0}^{\infty}$ is an increasing sequence of real numbers, $\{r_i\}_{i=0}^{n}$ so that for each r_i there as associated $F_{\leq r_i} C_k \subset C_k$ for every k that satisfies:

$$\{0\} \subset F_{\leq r_0} C_k \subset F_{\leq r_k} C_k \subset \cdots \subset F_{\leq r_n} C_k = C_k$$

As we well see, there are a lot of natural circumstances on which a filtration might arise. For example, the singular chain comples of a CW-complex is naturally filtered by its skeleton.
There are more structures we can put on a complex:

Remark 1. *A filtrated complex has a natural order on the generators compatible with the filtration.*

Definition 1.2 (Complex with ordered generators)**.** Let $\{C_k\}_{k=0}^{\infty}$ be a filtered chain complex, with some basis $\{e_j^{(i)}\}$. Then we say that $\{C_k\}_{k=0}^{\infty}$ has ordered generators when we fix the order $e_k^{(i)} < elj$ if $k = l$ and $i < j$.
Notice we do not compare generators that live on different chain groups.

Definition 1.3 (Canonical form)**.** Let $\{C_k\}_{k=0}^{\infty}$ be a chain complex, with some basis $\{e_k^{(i)}\}$. Then we say that $\{C_k\}_{k=0}^{\infty}$ is in canonical form if,

1. $\partial e_k^{(i)}$ is either 0 or another generator.
2. If $\partial e_k^{(i)} = \partial e_l^{(i)}$, then $e_k^{(i)} = e_l^{(i)}$.

Remark 2. *An equivalent formulation of the canonical form is that we can find a basis S of $\{C_k\}_{k=0}^{\infty}$ so that S can be separated into:*

1. *S_H: Generators of the homology of the complex.*
2. *S_{birth}: Births, that is, elements whose boundary is 0, but get killed in homology by an element of higher degree.*
3. *S_{death}: Deaths, elements whose boundary is another generator.*
4. *∂ is a bijection between S_{death} and S_{birth}.*

That is:

$$S = S_{birth} \sqcup S_{death} \sqcup S_H$$

Theorem 1. *[1] Every filtered chain complex can be reduced to one in canonical form by an upper-triangular change of basis which preserves the filtration.*
Furthermore, the canonical form is unique.

Corollary 1.1. *Two filtered complexes are equivalent if and only if they have the same canonical form.*

Corollary 1.2. *Category of filtered complexes is semi-simple. That is, any filtered complex, (or one with an ordered basis) can be expressed as a direct sum of 1 dimensional complexes with trivial differential and 2 dimensional complexes with trivial homology.*

Proof. Because any complex can we brought to canonical form, we can find a basis of the complex in the form of Remark 2. Then $\mathrm{Span}(S_H)$ splits as a sum of 1 dimensional complexes with trivial differential, since the differential is trivial on S_H. Similarly, the birth and death pairs splits into 2 dimensional complexes with one generator killing the other in homology.

QED

1.2 Augmented metric spaces

Definition 1.4. Let (X, d) be a metric space, we say it is augmented if it comes equiped with a function $f : X \to \mathbb{R}$.

We will say that an augmented metric space is injective if f is.
We will be interested in the case where X is finite.

1.3 Persistence complexes

Definition 1.5 (Persistent complex)**.** A persistence complex is a sequence of vector spaces $\{V_i\}_{0\leq i\leq n}$ and homomorphisms $f_i : V_i \to V_{i+1}$. That is, a diagram as follows:

$$V_0 \xrightarrow{f_0} V_1 \longrightarrow \cdots \xrightarrow{f_{n-1}} V_n$$

We denote $f_i^j = f_j \circ \cdots f_{i+1} \circ f_i$ whith $i < j$. Persistence diagrams may be of arbitrary cardinal, as long as the indexed is set is well-ordered.
But we shall only considered them when they are finite or countable.
We further define it's graded Betti number as:

$$b_i^j = \begin{cases} \dim(V_i) & \text{if} \quad i = j \\ \mathrm{rank}(f_i^j) & \text{if} \quad i < j \\ 0 & \text{if} \quad i > j \end{cases}$$

It is worth mentioning that we usually consider the maps f_i as inclusions and the spaces V_i as chain groups. As such, we can sometimes drop the vector space assumption and take modules over a PID. For example, it is common to consider $\mathbb{Z}$ instead of $\mathbb{K}$. We call this constructions persistent modules. Precisely we define:

Definition 1.6. (Persistence module) A persistence complex is a sequence of modules over a PID $\{M_i\}_{0\leq i\leq n}$ and homomophisms $f_i : M_i \to M_{i+1}$. With the rest of the definitions as above.

We give examples of useful persistence diagrams in 1.3.2.

Example 1.1 (Interval module)**.** *We define the interval module for $i < j$ to be the module where all maps are either the trivial ones, being the identity whenever possible and the zero map otherwise and the V_k to be:*

$$V_k = \begin{cases} \mathbb{K} & \textit{if } i \leq k \leq j \\ 0 & \textit{otherwise} \end{cases}$$

We denote it by $m(i, j)$ and we define it's associated interval to be $[i, j)$.

Proposition 1. *[2] The category of countable persistence complex is equivalent to the one of graded modules over* $\mathbb{K}[t]$

Proposition 2. *[2] Finite persistence diagrams are classified by their graded Betti numbers.*

Proposition 3. *[2] Every persistence module is the direct sum of interval modules. Furthermore, the multiplicity of each interval module is an invariant of the complex.*
This intervals are called the persistence intervals of a given persistence module.

One important observation about the previous proposition, is that a complex in canonical form gives a clear idea of what the persistence diagrams decomposition is. In particular, the beginning of a interval indicates a birth and the end of an interval indicates a death. And so, an element represent a non-trivial cohomology class if and only if it belongs to an interval of the form $[\alpha, \infty)$

1.3.1 Representations of the persistent complexes

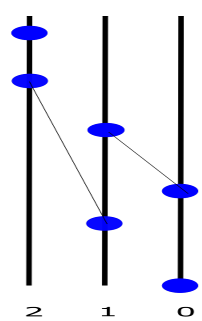
Firstofall, we can represent canonical form as in the following example:
Clearly, from any complex in canonical form we can construct an associated persistent diagram by reversing the above process. In the above example we ignored the filtration of the complex, but we consider that the diagram is filtered by height. Therefore:
Persistence diagrams are not the only way we can represent a canonical form. Another important example are codebars.

Definition 1.7 (Codebars)**.** By codebar we will mean a representation of similar to one below, that a collection of bars, each one index (although index might be repeated), where finite bars represent births and deaths and infinite bars represent elements that generate the homology of the complex.

Example 1.2. *Consider the following complex:*

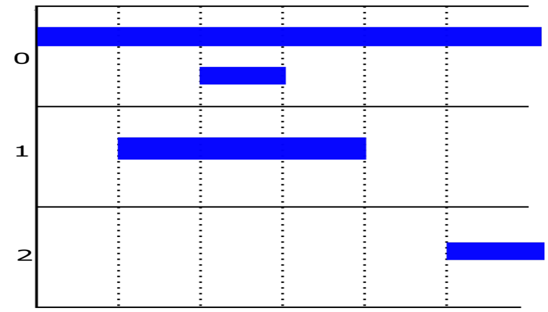
1. $C_0 = \{e_0^{(0)}, e_1^{(0)}\}$
2. $C_1 = \{e_0^{(1)}, e_1^{(1)}\}$
3. $C_0 = \{e_0^{(2)}, e_1^{(2)}\}$
4. $\partial e_1^{(1)} = e_0^{(1)}$, $\partial e_0^{(2)} = e_0^{(1)}$ *and* 0 *on the rest of the generators.*

And obviosuly $S_{birth} = \{e_0^{(1)}, e_0^{(1)}\}$, $S_{death} = \{e_1^{(1)}, e_0^{(2)}\}$, $S_H = \{e_0^{(0)}, e_1^{(2)}\}$
Then we have that it's persistence codebar would be:



And it's persistence diagram would be the following:

This example represents the barcode associated to the complex of the previous example.



1.3.2 Some important examples of persistent complexes

Definition 1.8 (Discrete morse complex)**.** Let $f : M \to \mathbb{R}$ be a morse function from a compact manifold with finitely many critical points $t_0 < \cdots < t_n$. Then the following diagram is a persistence complex:

$$H^{dR}(f^{-1}(-\infty, t_0)) \xrightarrow{\subseteq} H^{dR}(f^{-1}(-\infty, t_1)) \xrightarrow{\subseteq} \cdots \xrightarrow{\subseteq} H^{dR}(f^{-1}(-\infty, t_n))$$

Where H^{dR} is the de Rham complex and $\subseteq$ is homomorphism induced by the natural inclusion.

Definition 1.9 (Čech Complex)**.** Let $X \subset \mathbb{R}^n$ be a discrete subset. Then, for $d > 0$, we define the Čech complex of level ϵ as the following simplicial set:

$$\tilde{C}_\epsilon(X) = \{\sigma \subset X : \bigcap_{y \in \sigma} B(y, \epsilon) \neq \emptyset\}$$

The Čech complex is defined as $\{\tilde{C}_\epsilon\}_{\epsilon > 0}$

However, in practice the Čech complex can get too large to handle, notice that when ϵ is large enough, the Čech complex equals the power set of X. One solution to this is using the Vietories-Rips complex.

Definition 1.10 (Vietoris-Rips Complex)**.** Let $X \subset \mathbb{R}^n$ be a discrete subset. Then, for $d > 0$, we define the Čechcomplex of level ϵ as the following simplicial set:

$$\tilde{R}_\epsilon(X) = \{\sigma \subset X : \text{diam}(\sigma) \leq \epsilon\}$$

The Vietoris-Rips complex is defined as $\{\tilde{R}_\epsilon\}_{\epsilon > 0}$.

Remark 3. *Notice than if both the Čech and Vietories-Rips complex, the complex grows in a discrete manner meaning that there are only finitely many ϵ values for which the complex changes if ϵ is pertubed slightly.*
Furthermore, this complex are naturally filtrated and thus, they can be brought into canonical form.

There are many softwares that can be used to compute the Čech and Vietories-Rips complexes, most of them rely on bringing the complex to canonical form.
Another important example is the following:

Definition 1.11 (Bifiltrated Vietories-Rips complex)**.** Let $\mathcal{X} = (X, d, f)$ be an augmented metric space, let $X_\sigma = f^{-1}(-\infty, \sigma)$ we define the bifiltrated Vietories-Rips complex as the complex $\{\tilde{R}^\sigma_\epsilon\}_{\epsilon > 0, \sigma}$ with:

$$\tilde{R}^\sigma_\epsilon = \tilde{R}_\epsilon(X_\sigma)$$

1.4 Motivation

1.5 Example of canonical form

The proof of theorem 1 gives an algorithmic way to bring a complex into canonical form, we now expose how this methods works with an example. The general case should be clear afterwards.

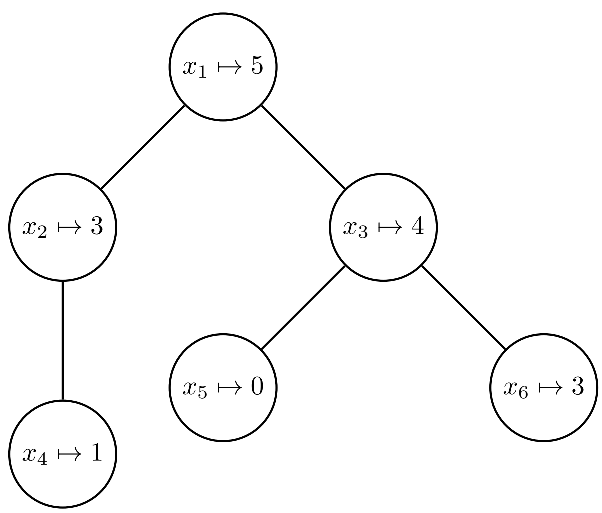
Elder-rule Staircodes

Morse functions on point clouds

We now focus on discrete augmented metric spaces. In particular, suppose we have a compact smooth manifold M equipped with a Morse function $f : M \to \mathbb{R}$.
From this manifold M we sample a finite amount of points $X = \{x_i\}_{i=0}^{n}$ and we record the value of f on X.
We investigate what sort of topological information we can recover from this discrete sample.
The first thing we would like to do is recover the gradient of the function. For this purpose we create a graph using X as its base set and we define a flow on the graph by moving from one point to the point connected with it where the morse function f takes the smallest value.

Definition 3.1 (NNG graph)**.** Given X a discrete set, we define $G^{-}(X)$ it's descending pseudo-gradient graph as the nearest neighbour graph of X.
That is the graph where the edges $\{x_i, x_j\}$ are precisely where either $d(x_i, x_j) = d(x_i, X)$ or $d(x_i, x_j) = d(x_j, X)$. Notice that the binary relation $x_i \sim x_j$ if and only if x_j is the closest vertex to x_i is not symmetric.
We denote as $\omega^{-} : G \to G$ the application that maps $p \in X$ to the result of iteratively following the point with that its joined to p and that minimises the value of f.
We define α^{-} in an analogous manner, by following the direction of maximum growth of f.

Example 3.1. *For example, let's consider the following graph, where we denote the value of f by $x_i : f(x_i)$:*



Then $\omega^{-}(x_1) = x_4$. So ω maps to local minimums, but not necessarily to absolute ones, even when restricting to the same component.

Bibliography

[1] Serguei Barannikov. The framed morse complex and its invariants. 1994.

[2] Afra Zomorodian Gunnar Carlson. Computing persistent homology. 2002.