Multiparameter persistence computation: review

Paul Snopov

June, 2022

In topological data analysis, one may use persistent homology to extract the topological features of the data. Its construction isn't too hard: one needs to build a simplicial complex K over the data, obtain a filtration of the complex (see fig. 1), that is a sequence of its subspaces $K_0 \subseteq K_1 \subseteq ... \subseteq K_n = K$, and then apply the i-th homology functor with coefficients in some field k to it, obtaining the next sequence of modules:

$$H_i(K_0) \to H_i(K_1) \to ... \to H_i(K).$$

In such setting, the algorithm for computing the persistent homology relies on Gaussian elimination and has its roots in the structure theorem for finitely generated modules over a PID. This is because the persistent module, as above, can be seen as a finitely generated module over $k[t]$. This gives the decomposition of persistent homology into the direct sum of indecomposable summands, which are usually called *interval modules*.

But in practice we often encounter richer structures that are described by multiple parameters. Such structures may be modeled with *multifiltrations*, like on the fig. 2. Such setting is much more complicated [2]. But a complete solution to the problem of computing multiparameter persistence exists and is provided by the Gröbner bases.

Let's give the precise definitions. We say that a topological space X is *multifiltered* if we're given a family of subspaces $\{X_v\}_{v\in\mathbb{N}^n}$ with inclusions $X_u \subseteq X_w$ whenever $u \leq w$ so that the diagrams

$$\begin{tikzcd} X_v \arrow[r] \arrow[d] & X_{v1} \arrow[d] \\ X_{v2} \arrow[r] & X_w \end{tikzcd}$$

commute. We will consider such multifiltered complexes, where each has has a unique minimal critical grade at which it enters the complex. Such multifiltrations are called *one-critical* and mostly arise in practice.

Given a simplicial complex K, we may define *chain groups* C_i as the free abelian groups on oriented i-simplices. We may provide the *boundary operator* $\delta_i C_i \to C_{i-1}$, that connects the chain groups into a *chain complex* $C_\bullet$:

$$... \to C_{i+1} \to C_i \to C_{i-1} \to ....$$

And for any chain complex, we may define the *ith homology group* H_i:

$$H_i(C_\bullet) \coloneqq \frac{\ker\, \delta_i}{\mathrm{im}\, \delta_{i+1}}.$$

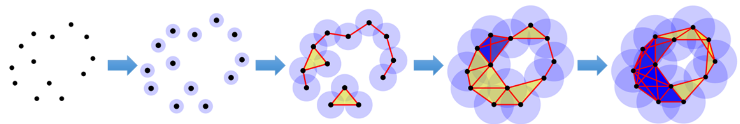

Figure 1 Example of a filtration of a simplicial complex built over the data [3]

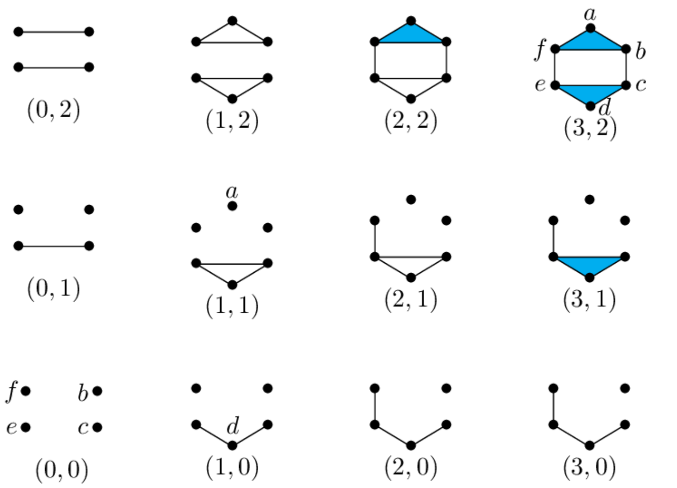

Figure 2 An example of a bifiltration of a complex at coordinate $(3, 2)$ [4]

Given a multifiltration $\{X_u\}_u$, $i : X_u \to X_v$ induces a map $i_* : H_i(X_u) \to H_i(X_v)$ at the homology level. The *ith persistent homology* H_i^{pers} is the image of i_* for all pairs $i \leq v$.

As was mentioned above, in the setting with a filtration, persistent homology corresponds to a graded $k[t]$-module. In the same way, persistent homology in the multifiltered setting corresponds to a finitely generated n-graded module over $k[t_1, ..., t_n]$. Moreover, the next theorem holds:

Theorem 1 (Realization [2])**.** *Let $k = \mathbb{F}_p$, $i \in \mathbb{N}$, M be an n-graded module over $k[t_1, ..., t_n]$. Then there's a multifiltered finite simplicial complex X so that $H_i^{pers}(X, k) \cong M$.*

So one can build a chain module over a multifiltered complex; let $\{K_u\}_u$ be a multifiltered simplicial complex, then the *ith chain module* is the n-graded module over $k[t_1, ..., t_n]$

$$C_i = \bigoplus_u C_i(K_u),$$

where the k-module structure is provided via the universal property of direct sum and $x^{v-u} : C_i(K_u) \to C_i(K_v)$ is the inclusion $K_u \to K_v$. For one-critical filtrations, these modules are free; the *standard basis* for the ith chain module C_i is given by the set of i-simplices in critical grades.

So, given standard bases, we may write the boundary operator $\delta_i : C_i \to C_{i-1}$ explicitly as a matrix with polynomial entries. This gives as new n-graded chain complex that encodes the multifiltration. The homology of this chain complex is precisely the persistent homology of the multifiltration. Then, by definition, homology can be computed in three steps:

1. Compute im δ_{i+1}: this problem is the *submodule membership problem*, which may be solved by computing the *reduced Gröbner bases* using the `Buchberger`, reduction and division algorithms.

2. Compute ker δ_i: the *(first) syzygy module* can be computer using `Schreyer's` algorithm

3. Compute H_i: one the above two tasks are complete, this is simple: we need to test whether the generators of the syzygy submodule are in the boundary submodule.

The submodule membership problem is a generalization of the *Polynomial Ideal Membership Problem*, which is `Exspace`-complete. But the multifiltrations provide the additonal structure that is used to simplify the algorithms; they key property is *homogeneity*: a matrix M with monomial entries is *homogeneous* if:

1. every column f of M is associated with a coordinate in the multifiltration u_f, and thus a corresponding monomial x^{u_f},

2. every non-zero element M_{jk} may be expressed as the quotient of the monomials associated with column k and row j. resp.

Any vector f endowed with a coordinate u_f that may be written as above is *homogeneous.*

With this in mind, one can simplify the algorithms [1]:

Lemma 1. *For a one-critical multifiltration, the matrix of $\delta_i : C_i \to C_{i-1}$ written in terms of the standard bases is homogeneous.*

Corollary 1. *For a one-critical multifiltration, the boundary matrix δ_i in terms of the standard bases has monomial entries.*

Lemma 2. *The S-polynomial $S(f, g)$ of homogeneous vectors f and g is homogeneous.*

Lemma 3. *The reminder of the division of homogeneous vector f by the tuple of the homogeneous vectors $(f_1, ..., f_t)$ is homogeneous.*

Theorem 2 (homogeneous Gröbner)**.** *The `Buchberger` algorithm computes a homogeneous Gröbner basis for a homogeneous matrix.*

Theorem 3 (homogeneous syzygy)**.** *For a homogeneous matrix, all matrices encountered in the computation of the syzygy module are homogeneous.*

Using optimization techniques (e.g. proper data structures), one can achieve the following result:

Theorem 4 ([1])**.** *Multidimensional persistence may be computed in polynomial time.*

References

[1] Carlsson, G., Singh, G., Zomorodian, A. (2009). Computing Multidimensional Persistence. In: Dong, Y., Du, DZ., Ibarra, O. (eds) Algorithms and Computation. ISAAC 2009. Lecture Notes in Computer Science, vol 5878. Springer, Berlin, Heidelberg.

[2] Carlsson, G., Zomorodian, A. The Theory of Multidimensional Persistence. Discrete Comput Geom 42, 71–93 (2009).

[3] Chi Seng Pun, Kelin Xia, Si Xian Lee Persistent-Homology-based Machine Learning and its Applications – A Survey.

[4] Heather A. Harrington, N. Otter, H. Schenck, U. Tillmann Stratifying Multiparameter Persistent Homology. SIAM J. Appl. Algebra Geom 3, 439-471 (2019)