AMAT 584 Lecture 37 4/27/20

Today: The persistence algorithm, continued

Review: To compute the persistent homology of a filtration F with largest simplicial complex F_{max}:

1) For each $j \geq 0$, order simplices of F_{max} with respect to their birth time in the filtration.

2) Using these orderings, construct a "boundary matrix" D containing all the matrices $[\delta_j]$ as blocks just above the diagonal.

Columns and rows of D correspond to simplices of F_{max}.

3) Use left-to-right column additions to transform D into a reduced matrix R.

4) Read each barcode off of R, using the following formula:

$Barc(H_i(F)) =$
$\{[birth(\sigma), birth(\tau)) \mid \text{pivot of column } \tau \text{ in } R \text{ is } \sigma \text{ and } dim(\sigma) = i\}$

$\{[\text{birth}(\sigma), \infty) \mid \text{col } \sigma = 0, \sigma$ is not the pivot of any column in R, $\dim(\sigma) = i\}$.

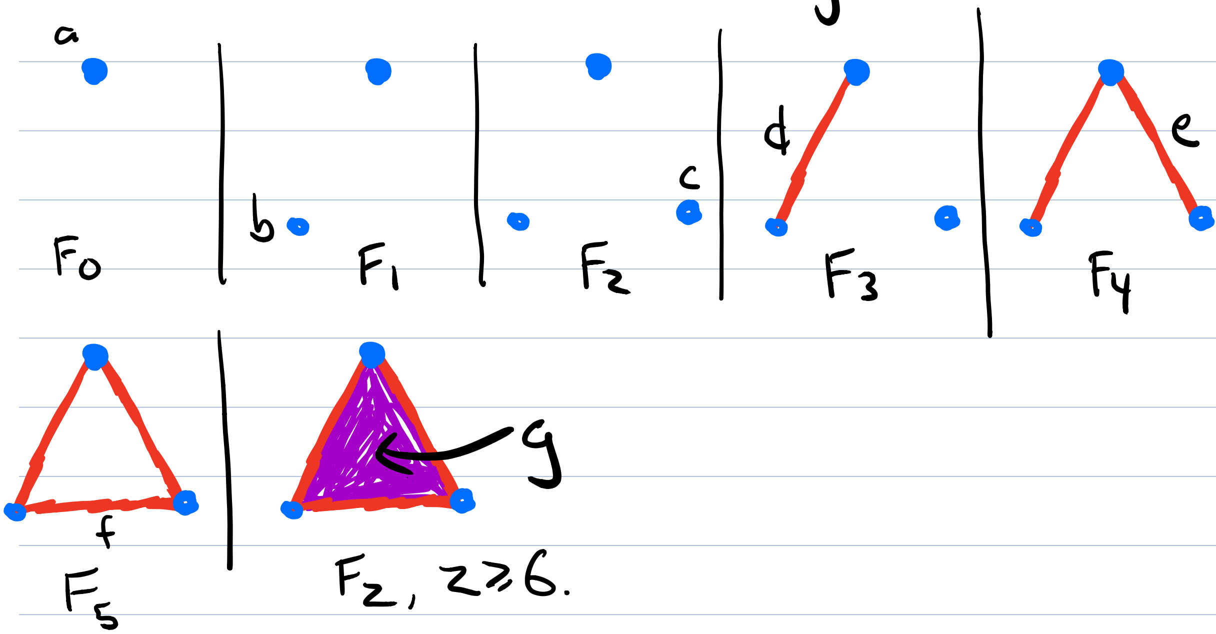
Note: Any interval the form $[z, z)$ is ignored.

In words: Every non-zero column in R gives a finite-length interval in one of the barcodes, by pairing the column with its pivot.

A zero column which doesn't appear as the pivot of any other column gives an infinite interval.

<u>Example of the last step</u>

Last lecture we considered the following filtration F:



We saw that for this filtration,

$$D = \begin{array}{c} \\ a\\ b\\ c\\ d\\ e\\ f\\ g \end{array} \begin{array}{c} \begin{array}{ccccccc} a & b & c & d & e & f & g \end{array} \\ \left(\begin{array}{ccc|ccc|c} & & & 1 & 1 & 0 & \\ & 0 & & 1 & 0 & 1 & 0 \\ & & & 0 & 1 & 1 & \\ \hline & & & & & & 1 \\ & 0 & & & 0 & & 1 \\ & & & & & & 1 \\ \hline & 0 & & & 0 & & 0 \end{array} \right) \end{array}$$

Applying the standard reduction (Step 3) gives

$$R \quad \begin{array}{r} \\ F^0_{max} \left\{ \begin{array}{c} a\\ b\\ c \end{array} \right. \\ F^1_{max} \left\{ \begin{array}{c} d\\ e\\ f \end{array} \right. \\ F^2_{max} \left\{ \begin{array}{c} g \end{array} \right. \end{array} \begin{array}{c} \begin{array}{ccccccc} a & b & c & d & e & f & g \end{array} \\ \left(\begin{array}{ccc|ccc|c} & & & 1 & 1 & 0 & \\ & 0 & & 1 & 0 & 0 & 0 \\ & & & 0 & 1 & 0 & \\ \hline & & & & & & 1 \\ & 0 & & & 0 & & 1 \\ & & & & & & 1 \\ \hline & 0 & & & 0 & & 0 \end{array} \right) \end{array}$$

(The only column of D that was changed was column f).

We get a finite interval in one of the barcodes for each non-zero column:

$$\underbrace{[\mathrm{birth}(\overset{F^0_{max}}{\overset{\cup}{b}}), \mathrm{birth}(d))}_{=[1,3)}, \underbrace{[\mathrm{birth}(\overset{F^0_{max}}{\overset{\cup}{c}}), \mathrm{birth}(e))}_{=[2,4)}, \underbrace{[\mathrm{birth}(\overset{F^1_{max}}{\overset{\cup}{f}}), \mathrm{birth}(g))}_{=[5,6)}$$

The zero columns are a, b, c, f. Of these, all but a is a pivot of another column.

So the only infinite interval is $\underbrace{[\mathrm{birth}(\overset{F^0_{max}}{\overset{\cup}{a}}), \infty)}_{=[0,\infty)}$

To figure out which interval goes in which barcode, we look at the dimension of the simplex ending the left endpoint.

$$\mathrm{Barc}(H_0(F)) = \{[1,3), [2,4), [0,\infty)\}$$
$$\mathrm{Barc}(H_1(F)) = \{[5,6)\}.$$

Note: If we only want to compute persistent homology up to degree k, i.e. $\mathrm{Barc}(H_i(F))$ for $i \in \{0,1,\dots,k\}$, then we don't need to consider all of F. It's sufficient to consider the subfiltration consisting of only simplices of dimension $\leq k+1$.

This often allows us to keep the size of D much smaller

An example coming from data

Let $X = \{\overset{a}{\overset{\shortparallel}{(0,0)}}, \overset{b}{\overset{\shortparallel}{(2,0)}}, \overset{c}{\overset{\shortparallel}{(0,2)}}, \overset{d}{\overset{\shortparallel}{(2,2)}}\} \subset \mathbb{R}^2$, given the Euclidean metric.

Let $B_i = \mathrm{Barc}(H_i(\mathrm{Rips}(X)))$

We will compute each B_i.

Note that $\mathrm{Rips}(X)$ is a $[0,\infty)$-index filtration. We will show by example that the persistence algorithm works in essentially the same way for such a filtration as in the $\mathbb{N}$-indexed case.

Let $F = \mathrm{Rips}(X)$. F_{max} is the 3-simplex with vertex set X.

Let's give names to all the simplices of F_{max}, and record their birth times:

to keep notation simple, we write $\mathrm{birth}(\sigma)$ as $\beta(\sigma)$.

$a = [a]$, $b = [b]$, $c = [c]$, $d = [d]$

$\beta(a) = \beta(b) = \beta(c) = \beta(d) = 0$

$e=[a,b]$, $f=[b,c]$, $g=[c,d]$, $h=[d,a]$, $i=[a,c]$, $j=[b,d]$
$\beta(e)=\beta(f)=\beta(g)=\beta(h)=2.$
$\beta(i)=\beta(j)=2\sqrt{2}.$

$k=[a,b,c]$, $l=[a,b,d]$, $m=[a,c,d]$, $n=[b,c,d]$

$\beta(k)=\beta(l)=\beta(m)=\beta(n)=2\sqrt{2}$

$o=[a,b,c,d]$

$\beta(o)=2\sqrt{2}.$

Note: The alphabetical order on each X^i is in order of increasing birth

to be continued: