

VI.2 Stability

Persistent homology is stable, that is, small perturbations of the input only cause small perturbations of the output. This property has a number of non-trivial consequences, some of which will be discussed in Section VI.3. In this section, we study how changing the filtration affects the corresponding reduced matrix and use the gained insights to prove the stability of persistence diagrams.

Matrix reduction revisited. We recall the notion of a reduced 0-1 matrix in which each non-zero column has its lowest one in a unique row. Starting with the boundary matrix, D , whose rows and columns correspond to the simplices in the order they enter the filtration, we use the algorithm of the previous section to reduce the matrix by adding columns from the left. The resulting reduced matrix can therefore be written as $R = DV$, where V is an invertible, upper triangular 0-1 matrix. Let U be the right inverse of V and note that it is again invertible and upper triangular. By multiplication from the right we get $RU = DVU = D$ which we call an *RU-decomposition* of the boundary matrix. It is characterized by R being reduced and U being upper triangular. Not surprisingly, the RU-decomposition is not unique. For example, we could use additional column operations to remove as many ones from the matrix as we can. On the other hand, the pairing implied by the lowest ones in the columns is unique. To prepare the proof of this claim we write $R^{i,j}$ for the lower left submatrix obtained by deleting the first $i - 1$ rows and the last $n - j$ columns from R , as illustrated in Figure VI.6. Any linear combination of non-zero

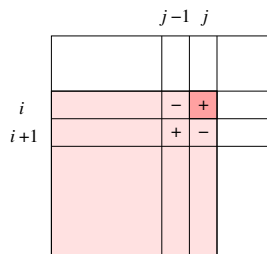


Figure VI.6: The shaded submatrix $R^{i,j}$ of R . We have $i = \text{low}(j)$ iff the indicated inclusion-exclusion formula of ranks of lower left submatrices gives 1.

columns in $R^{i,j}$ has its lowest one in the same row as the lowest of the lowest ones of the involved columns. The combination is non-zero implying that the combined columns are linearly independent. In other words, the rank of $R^{i,j}$

is equal to its number of non-zero columns. We consider the same collection of submatrices of the boundary matrix, and for each index pair (i, j) we define

$$r_D(i, j) = \text{rank } D^{i,j} - \text{rank } D^{i+1,j} + \text{rank } D^{i+1,j-1} - \text{rank } D^{i,j-1}.$$

We prove shortly that the pairing function can be expressed in terms of r_D and is therefore independent of the particular reduced matrix we derive from D .

PAIRING LEMMA. Letting $D = RU$ be an RU-decomposition, we have $i = \text{low}(j)$ iff $r_D(i, j) = 1$. In particular, the pairing defined by the lowest ones in the reduced matrix does not depend on R .

PROOF. We have $r_D = r_R$ since adding columns from the left maintains the rank of every lower left submatrix. It therefore suffices to prove that $i = \text{low}(j)$ iff $r_R(i, j) = 1$. First assume $i = \text{low}(j)$. We recall that the rank of $R^{i,j}$ is equal to its number of non-zero columns. The last column is non-zero, so $\text{rank } R^{i,j} - \text{rank } R^{i,j-1} = 1$. Deleting the top row makes the last column zero, so $\text{rank } R^{i+1,j} - \text{rank } R^{i+1,j-1} = 0$, as required. Second assume $i \neq \text{low}(j)$. If $\text{low}(j) < i$ then the last column is zero, so $\text{rank } R^{i,j} - \text{rank } R^{i,j-1} = 0$ and $\text{rank } R^{i+1,j} - \text{rank } R^{i+1,j-1} = 0$. If $\text{low}(j) > i$ then the last column is non-zero even after deleting the first row, so $\text{rank } R^{i,j} - \text{rank } R^{i,j-1} = 1$ and $\text{rank } R^{i+1,j} - \text{rank } R^{i+1,j-1} = 1$. In both cases the claim follows. \square

Transpositions. Suppose we change the order of just two simplices, transposing the two corresponding rows and columns in the boundary matrix. The new boundary matrix is PDP , where P is the permutation matrix that swaps i with $i + 1$. Multiplication with P from the left swaps the two rows and multiplication from the right swaps the two columns. Letting $D = RU$ be an RU-decomposition we now get $PRUP = (PRP)(PUP)$, but this is not necessarily an RU-decomposition of PDP . It can fail to be one because PRP is not reduced or because PUP is not upper triangular. However, in each case this can be remedied with relatively little effort. Recall that a simplex is positive if its addition to the complex increases the Betti number of the same dimension. It corresponds to a zero column in R . A simplex is negative if its addition to the complex decreases the Betti number of one lower dimension. It corresponds to a non-zero column with a lowest one in R .

Case 1. The simplices in positions i and $i + 1$ are both positive. Then column i is zero so we may set $U[i, i + 1] = 0$. It follows that PUP is upper triangular and we only need to worry about PRP . If it fails to be reduced, as in Figure VI.7 on the left, we add column k to column l and thus obtain an RU-decomposition.

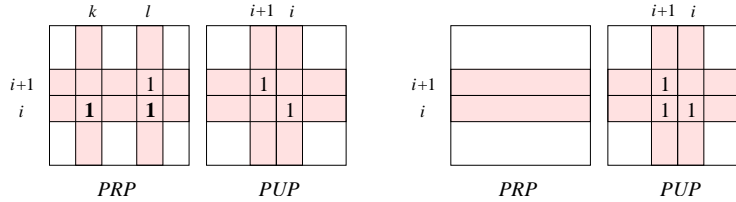


Figure VI.7: Left: after swapping the two positive simplices at positions i and $i + 1$ the matrix on the left may no longer be reduced. Right: after swapping the two negative simplices at positions i and $i + 1$ the matrix on the right may no longer be upper triangular.

Case 2. The simplices in positions i and $i + 1$ are both negative. The corresponding rows do not contain any lowest ones so we only need to worry about *PUP*. It fails to be upper triangular iff $U[i, i + 1] = 1$, as in Figure VI.7 on the right. We fix the trouble by adding row $i + 1$ to row i in U and adding column i to column $i + 1$ in R . This does not affect the product of the two matrices. If $low(i) < low(i + 1)$ before this operation then the lowest ones remain unique and we have an RU-decomposition after the transposition. On the other hand, if $low(i) > low(i + 1)$ then we need to make the lowest ones unique again, which we do by adding column $i + 1$ to column i . After the transposition this is adding column i to column $i + 1$ and we get again an RU-decomposition.

Case 3. The simplex at position i is negative and that at position $i + 1$ is positive. Since row $i + 1$ has no lowest one we only need to worry about *PUP*. The only troublesome case is $U[i, i + 1] = 1$ and we can remedy the situation the same way as in Case 2.

Case 4. The simplex at position i is positive and that at position $i + 1$ is negative. Because row i has no lowest one *PRP* is reduced. Furthermore, column i is zero so we may set $U[i, i + 1] = 0$ to make sure that *PUP* is upper triangular.

The above algorithm maintains the RU-decomposition of the boundary matrix in a constant number of row and column operations. It thus takes linear time to perform a transposition of two contiguous simplices in the ordering of the filtration.

Switches. A transposition of two simplices may or may not change the pairing between positive and negative simplices. We call a change of the pairing a

switch since it occurs only locally, exchanging the simplices between two pairs. Switches are indicative of major topological changes during the rearrangement of the simplices in the filtration. Figure VI.8 illustrates the three types we observe, between two positive simplices, between two negative simplices, and between a positive and a negative simplex. In each case the two simplices have the same dimension.

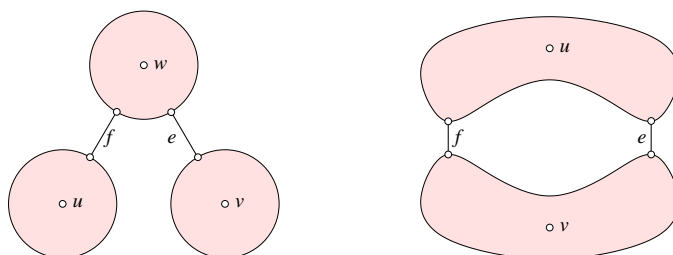


Figure VI.8: The vertices u , v , and w are the oldest in their respective components, which are eventually joined by the edges e and f .

Type I. Two positive simplices swap places and switch partners in the pairing.

This happens in Case 1 when the configuration is as shown in Figure VI.7. To construct an example we start with the left complex in Figure VI.8. The oldest vertices of the three components, u, v, w , enter the complex in this order, and the edges e and f eventually join the components, also in this order. We have (w, e) in the pairing because w is the younger of the oldest vertices of the two components joined by e . Similarly, we have (v, f) in the pairing. After swapping v and w we have (v, e) and (w, f) in the pairing.

Type II. Two negative simplices swap places and switch partners in the pairing. This happens in Case 2 when $low(i) > low(i + 1)$. To construct an example we start again with the left complex in Figure VI.8 assuming u, v, w, e, f enter the complex in this order. We now swap the two edges causing a switch from $(w, e), (v, f)$ to $(w, f), (v, e)$ in the pairing.

Type III. A positive simplex swaps its place with a negative simplex. This happens in Case 3 when $low(i) > low(i + 1)$, as for Type II. We construct an example from the right complex in Figure VI.8. Let u and v be the oldest vertices in their components, which are joined first by e and then by f . Suppose the hole created by f later fills up and t is the triangle that kills the corresponding homology class. We thus have (v, e) and (f, t)

in the pairing. After swapping the two edges we have (v, f) and (e, t) . Besides exchanging the two edges the switch also exchanges their types.

An interesting feature of all three types of switches is that they occur only between disjoint or nested pairs. As illustrated in Figure VI.9, the switch then maintains this property.

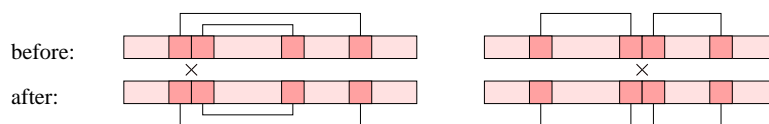


Figure VI.9: The ordered sequences of simplices before and after the switch. A Type I switch on the left and a Type III switch on the right. The Type II switch is symmetric to the Type I switch and obtained by reflecting the picture along a vertical line.

Bottleneck distance. The above analysis of transpositions and the related classification of switches will be instrumental in proving that persistence is stable. Another crucial ingredient is the appropriate notion of distance between two persistence diagrams, Dgm and Dgm' . Each diagram is a finite set of off-diagonal points in \mathbb{R}^2 . To each set we add an arbitrary, finite number of points on the diagonal such that the total number of points is the same in both sets. Then we form a perfect matching between the two sets and record the length of the longest edge measured as L_∞ -distance between its endpoints. The *bottleneck distance* between the two diagrams is the infimum over all choices of diagonal points, Δ and Δ' , and all perfect matchings,

$$d_B(\text{Dgm}, \text{Dgm}') = \inf_{\gamma} \max_u \|u - \gamma(u)\|_\infty,$$

where $\gamma : \text{Dgm} \cup \Delta \rightarrow \text{Dgm}' \cup \Delta'$ is a bijection between the two sets and $u \in \text{Dgm} \cup \Delta$ is a point in the first set. Letting Dgm'' be a third persistence diagram it is clear that $d_B(\text{Dgm}, \text{Dgm}'') \leq d_B(\text{Dgm}, \text{Dgm}') + d_B(\text{Dgm}', \text{Dgm}'')$. In words, the bottleneck distance satisfies the triangle inequality. It also satisfies $d_B(\text{Dgm}, \text{Dgm}') = 0$ iff $\text{Dgm} = \text{Dgm}'$, so it is a metric.

Stability of persistence. To formulate the result, we introduce a function $f : K \rightarrow \mathbb{R}$ that maps each simplex to a real number. We require that f be *monotone* by which we mean that $f(\tau) \leq f(\sigma)$ if τ is a face of σ . We obtain a filtration by ordering the simplices consistent with the function, that

is, τ precedes σ if $f(\tau) < f(\sigma)$. After pairing the simplices as before, we map each pair (τ, σ) to the point $(f(\tau), f(\sigma))$ in the plane. This gives a persistence diagram for each dimension p , which we denote as $\text{Dgm}_p(f)$. As long as the pairing does not change, small changes in the function correspond to small changes in the diagram. To prove stability we have to establish the same even if the pairing changes. Letting $g : K \rightarrow \mathbb{R}$ be a second monotone function we write $\|f - g\|_\infty = \max_{\sigma \in K} |f(\sigma) - g(\sigma)|$ for the largest absolute difference.

STABILITY THEOREM FOR FILTRATIONS. For any monotone functions $f, g : K \rightarrow \mathbb{R}$ and dimension p , the bottleneck distance between the two dimension p persistence diagrams is $d_B(\text{Dgm}_p(f), \text{Dgm}_p(g)) \leq \|f - g\|_\infty$.

PROOF. We transform f into g using the straight-line homotopy $f_\lambda : K \rightarrow \mathbb{R}$ defined by $f_\lambda(\sigma) = (1 - \lambda)f(\sigma) + \lambda g(\sigma)$ for $0 \leq \lambda \leq 1$. Assuming $f = f_0$ and $g = f_1$ are both injective, there are only finitely many values of λ for which the mixed function is not injective. The transpositions happen exactly at these values, which we denote as $\lambda_1 < \lambda_2 < \dots < \lambda_m$. It is convenient to add $\lambda_0 = 0$ and $\lambda_{m+1} = 1$ to the sequence.

We first consider two parameter values without transpositions between them, $\lambda_i < r < s < \lambda_{i+1}$. The lowest one pairing is the same for r and for s . Letting (τ, σ) be a member of the pairing, we have $u_r = (f_r(\tau), f_r(\sigma))$ in the diagram of f_r and $u_s = (f_s(\tau), f_s(\sigma))$ in the diagram of f_s . The L_∞ -distance between these two points is the larger of the two coordinate differences, which implies

$$\begin{aligned} d_B(\text{Dgm}_p(f_r), \text{Dgm}_p(f_s)) &\leq \|f_r - f_s\|_\infty \\ &= (s - r)\|f - g\|_\infty. \end{aligned}$$

A transposition changes the matching of simplices but it does not affect the persistence diagram. Hence,

$$\begin{aligned} d_B(\text{Dgm}_p(f), \text{Dgm}_p(g)) &\leq \sum_{i=0}^m d_B(\text{Dgm}_p(f_{\lambda_i}), \text{Dgm}_p(f_{\lambda_{i+1}})) \\ &\leq \sum_{i=0}^m (\lambda_{i+1} - \lambda_i)\|f - g\|_\infty. \end{aligned}$$

The sum of the $\lambda_{i+1} - \lambda_i$ is equal to 1 which implies the claimed inequality. \square

Bibliographic notes. The first proof of the stability of persistence diagrams has been given for functions; this will be described in the next section. Independently, d'Amico, Frosini and Landis proved the stability of dimension 0

persistence diagrams [2]. The proof of stability for filtrations presented in this section is taken from [1]. It has the benefit of being elementary and implying a linear time algorithm for maintaining an RU-decomposition of the boundary matrix under transpositions of simplices.

- [1] D. COHEN-STEINER, H. EDELSBRUNNER AND D. MOROZOV. Vines and vineyards by updating persistence in linear time. *In* “Proc. 22nd Ann. Sympos. Comput. Geom., 2006”, 119–126.
- [2] M. D’AMICO, P. FROSINI AND C. LANDI. Optimal matching between reduced size functions. Report 35, DISMI, Università di Modena e Reggio Emilia, Bologna, Italy, 2003.