Étude de couvertures de réseaux de métros, application de l'homologie persistante et optimisation.

Elowan ; 10381

May 13, 2025

Le déroulement

Trouver les zones les moins biens desservies par un réseau de métros.

Les principales étapes :

- Convertir les données géographique en espace métrique
- Créer une suite de formes géométriques à partir de cet espace
- Trouver les trous de ces formes

Plus en détail, reconnaitre un trou : l'idée du lasso

Chercher à enrouler au lasso, puis reduire ces lassos sans jamais le rompre ou déchirer la forme.

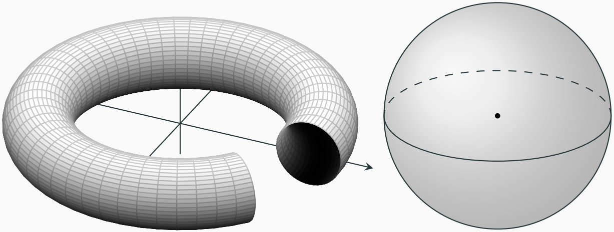


Figure 1: Un tore contenant deux trous et une sphère contenant 0 trou.

Définitions

Simplexe
Un simplexe σ de dimension k (ou k-simplexe) correspond à l'enveloppe convexe de $k+1$ vecteurs non inclus dans un sous-espace affine de dimension $k-1$.

Face
On dit que σ_i est une face de σ_j si et seulement si $\sigma_i \subset \sigma_j$ et la dimension de σ_i $\dim(\sigma_i)$ est égale à $\dim(\sigma_j) - 1$.

Complexe simplicial
Un ensemble de simplexes.

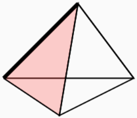


Figure 2: Exemple de complexe simplicial

Filtration
Suite croissante pour l'inclusion de complexes simpliciaux.

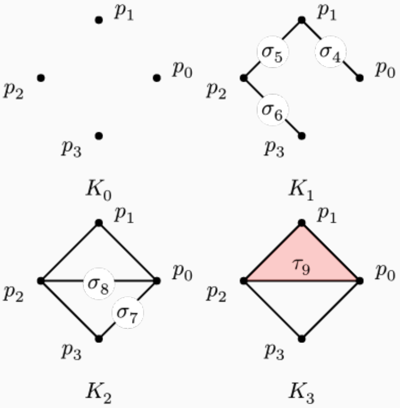


Figure 3: Exemple de filtration.

Classe d'homologie
Elle représente un trou en dimension n.

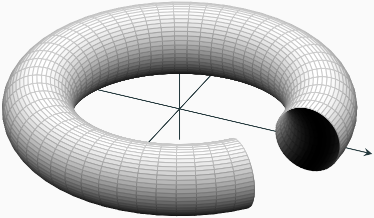


Figure 4: Le tore possède deux classes d'homologies, celle du trou au centre et celle formée par le cylindre.

La méthode de l'homologie persistante

- Construire une filtration à partir d'un ensemble de points
- Application de l'algorithme *standard*
- Récupération des classes d'homologies

Définition de la distance

Distance
On définit la distance d entre deux stations de metro x et y :

$$d(x, y) = \frac{1}{2}(min(t_{pied}(x, y), t_{voit}(x, y)) + min(t_{pied}(y, x), t_{voit}(y, x)))$$

Poids d'une station
La moyenne sur une semaine du temps d'attente en station.

Définition des complexes pondérés de Vietoris-Rips

Complexe Simplicial pondéré de Vietoris-Rips
On le définit au rang r, comme l'ensemble des simplexes $(x_{i_0}, ..., x_{i_k})$ tels que :

$$\begin{cases} \forall j \in [|0, k|], poids_{i_j} < r \\ \forall (j, l) \in [|0, k|]^2, d(x_{i_j}, x_{i_l}) + poids_{i_j} + poids_{i_l} < 2r \end{cases}$$

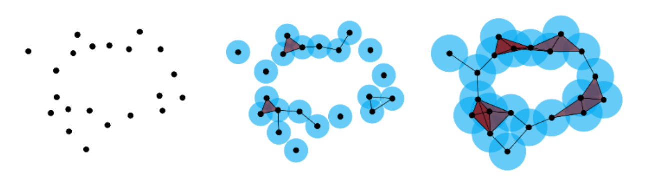


Figure 5: Image repr les vietoris rips, tiré de *Persitent homology for resource coverage: A case study of access to polling sites*

Préparatif de l'algorithme (1/2): Ordre total sur les simplexes

Soient une filtration $K_0 \subset K_1 \subset ... \subset K_p$ et l'ensemble S de tous les simplexes apparaissant dans la filtration. On indice S de sorte que pour tout σ_i et σ_j de S:

$$\left.\begin{array}{l}\text{Si } \sigma_i \in K_{k_i} \text{ et } \sigma_j \in K_{k_j} \text{ avec } k_i < k_j \\ \text{Sinon si } \sigma_i \text{ est une face de } \sigma_j\end{array}\right\} \Rightarrow i < j$$

L'ordre total est déduit de cet indiçage.

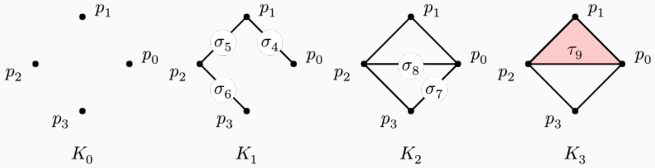
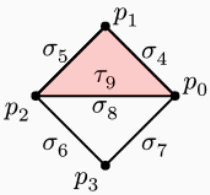


Figure 6: Indiçage de la filtration

Préparatif de l'algorithme (2/2): Matrice de brodure



	4	5	6	7	8	9
0	1			1	1	
1	1	1				
2		1	1		1	
3			1	1		
4						1
5						1
6						
7						
8						1
low	**1**	**2**	**3**	**3**	**2**	**8**

Figure 7: Exemple de matrice de bordure B associé à un ordre total

Application de l'algorithme

Algorithme Standard($B \in M_n(\{0,1\})$**)**

```
for j allant de 0 a n-1:
    while (il existe i < j avec low[i] = low[j]):
        ajouter colonne i de B a colonne j modulo 2
```

Compréhension du résultat en sortie

	4	5	6	7	8	9
0	1			1	1	
1	1	1				
2		1	1		1	
3			1	1		
4						1
5						1
6						
7						
8						1
low	**1**	**2**	**3**	**3**	**2**	**8**

	4	5	6	7	8	9
0	1					
1	1	1				
2		1	1			
3			1			
4						1
5						1
6						
7						
8						1
low	**1**	**2**	**3**	**-1**	**-1**	**8**

Figure 8: Matrice B et réduite $\overline{B}$

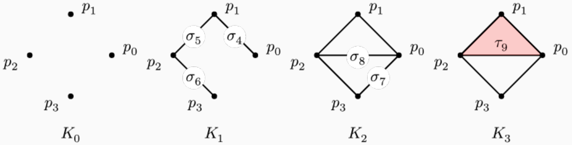


Figure 9: Filtration et utilisée

Optimisation

Remarques à l'initiative de la recherche d'optimisation :

- La matrice est creuse ;
- L'opération de somme de colonnes laisse invariante tous les simplexes de dimension différente.

Modifications apportées :

- Représentation de la matrice en liste d'adjacence (double liste chaînée ordonnée) ;
- Application de l'algorithme sur des matrices extraites et non la totale.

Algorithme Standard optimisé($B \in M_n(\{0,1\})$)

```
dims <- Tableau des simplexes ou dims[i] contient la
    liste des simplexes de dimension i
for toute dimension d a considerer:
    for chaque simplexe j de dims[d]
        while (il existe i dans dims[d] tel que low[j]
            = low[i]):
            ajouter colonne i de B a colonne j modulo
                2
```

Résultats

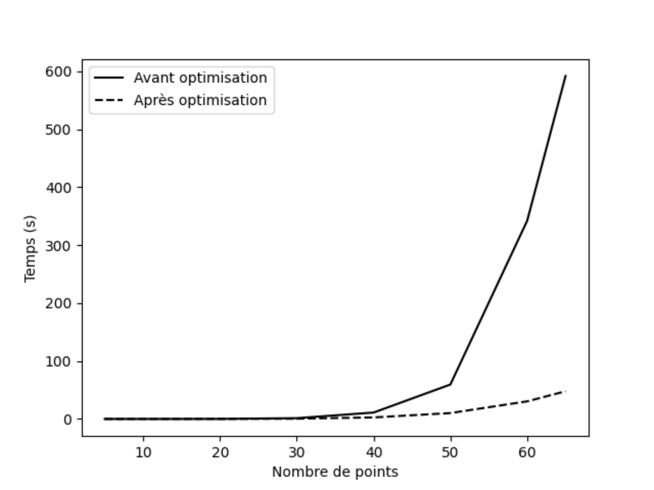


Figure 10: Ordonnée linéaire

Résultats

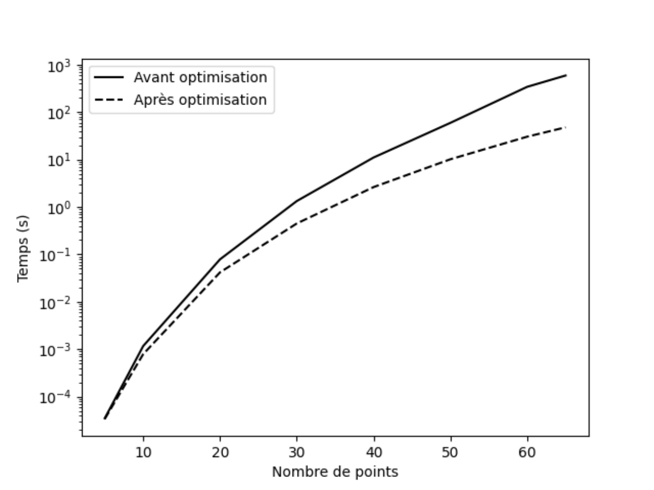


Figure 11: Ordonnée logarithmique

Résultats

Figure 12: Marseille

Figure 13: Toulouse

Annexe : Définition

Définition d'un cycle
Un cycle est une sous-variété fermée.

Définition d'une limite
Une limite est un cycle qui est également la limite d'une sous-variété

Définition d'une classe d'homologie
Une classe d'homologie est une classe d'équivalence de cycles modulo une limite : elle est donc représentée par un cycle qui n'est la limite d'aucune sous-variété, il représente donc un trou, une variété dont la limite serait ce cycle, mais qui n'est pas là

Annexe : Diagrammes de persistance

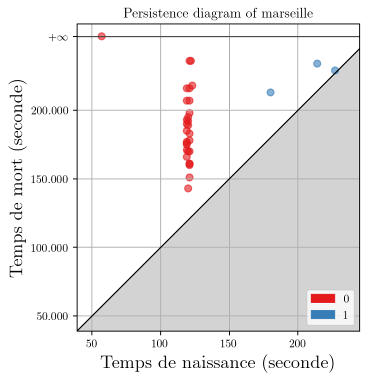


Figure 14: Diagramme de persistance de Marseille

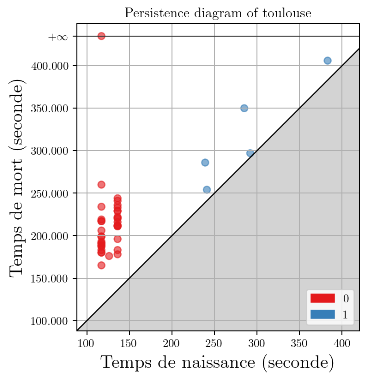


Figure 15: Diagramme de persistance de Toulouse

Annexe : Récupération des données

Pour le calcul des temps de trajet : apidocs.geoapify.com

Pour la récupération des stations et des temps d'attentes moyens : transport.data.gouv.fr