

Introduction à la théorie de la Persistance

Francis Lazarus

GIPSA-Lab, CNRS

19 Mars 2010

Coder les variations topologiques

Partant d'une filtration d'un espace X :

$$X_1 \subset X_2 \subset \dots \subset X_n = X$$

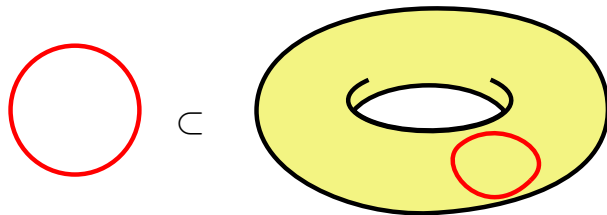
On en déduit :

$$H_*(X_1) \rightarrow \dots \rightarrow H_*(X_n)$$

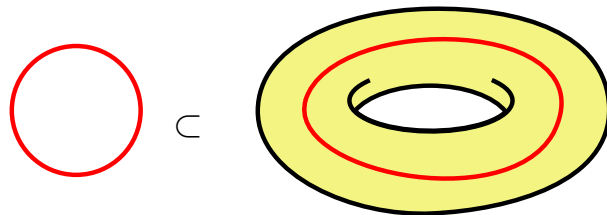
C'est un objet algébrique que l'on va coder par des intervalles.

- 1 Dans la pratique on se donne $f : X \rightarrow \mathbb{R}$ et $X_i = f^{-1}([-\infty, a_i])$
- 2 On peut remplacer les \subset par des applications continues.
- 3 Dans la pratique on travaille avec des complexes simpliciaux.

Exemple



$$\mathbb{K} \xrightarrow{0} \mathbb{K}^2$$



$$\mathbb{K} \xrightarrow{Id \times 0} \mathbb{K}^2$$

Théorie de la persistance homologique

- Robins, V. **Towards computing homology from finite approximations** *Topology proceedings*, 1999, V. 24.
- Edelsbrunner, H., Letscher, D. and Zomorodian, A. **Topological Persistence and Simplification**. *IEEE Symposium on Foundations of Computer Science*, 2000.
- Zomorodian, A. and Carlsson, G. **Computing Persistent Homology**. *Discrete & Computational Geometry*, 2005, V. 33.

Homologie simpliciale

- K : collection de simplexes orientés contenant leurs faces.
- $C_*(K)$: combinaisons formelles sur \mathbb{K} de simplexes.
- $\partial : C_*(K) \rightarrow C_*(K)$: opérateur bord.

$$\partial[s_0, s_1, \dots, s_n] = \sum_{i=0}^n (-1)^i [s_0, \dots, \hat{s}_i, \dots, s_n]$$

$$\partial \circ \partial = 0$$

Donc $B_*(K) \subset Z_*(K)$ où $B_*(K) = \text{Im } \partial$ et $Z_*(K) = \ker \partial$.

$$H_*(K) = Z_*(K)/B_*(K)$$

Homologie simpliciale

- K : collection de simplexes orientés contenant leurs faces.
- $C_*(K)$: combinaisons formelles sur \mathbb{K} de simplexes.
- $\partial : C_*(K) \rightarrow C_*(K)$: opérateur bord.

$$\partial[s_0, s_1, \dots, s_n] = \sum_{i=0}^n (-1)^i [s_0, \dots, \hat{s}_i, \dots, s_n]$$

$$\partial \circ \partial = 0$$

Donc $B_*(K) \subset Z_*(K)$ où $B_*(K) = \text{Im } \partial$ et $Z_*(K) = \ker \partial$.

$$H_*(K) = Z_*(K)/B_*(K)$$

Homologie simpliciale

- K : collection de simplexes orientés contenant leurs faces.
- $C_*(K)$: combinaisons formelles sur \mathbb{K} de simplexes.
- $\partial : C_*(K) \rightarrow C_*(K)$: opérateur bord.

$$\partial[s_0, s_1, \dots, s_n] = \sum_{i=0}^n (-1)^i [s_0, \dots, \hat{s}_i, \dots, s_n]$$

$$\partial \circ \partial = 0$$

Donc $B_*(K) \subset Z_*(K)$ où $B_*(K) = \text{Im } \partial$ et $Z_*(K) = \ker \partial$.

$$H_*(K) = Z_*(K)/B_*(K)$$

Homologie simpliciale

- K : collection de simplexes orientés contenant leurs faces.
- $C_*(K)$: combinaisons formelles sur \mathbb{K} de simplexes.
- $\partial : C_*(K) \rightarrow C_*(K)$: opérateur bord.

$$\partial[s_0, s_1, \dots, s_n] = \sum_{i=0}^n (-1)^i [s_0, \dots, \hat{s}_i, \dots, s_n]$$

$$\partial \circ \partial = 0$$

Donc $B_*(K) \subset Z_*(K)$ où $B_*(K) = \text{Im } \partial$ et $Z_*(K) = \ker \partial$.

$$H_*(K) = Z_*(K) / B_*(K)$$

Homologie simpliciale

- K : collection de simplexes orientés contenant leurs faces.
- $C_*(K)$: combinaisons formelles sur \mathbb{K} de simplexes.
- $\partial : C_*(K) \rightarrow C_*(K)$: opérateur bord.

$$\partial[s_0, s_1, \dots, s_n] = \sum_{i=0}^n (-1)^i [s_0, \dots, \hat{s}_i, \dots, s_n]$$

$$\partial \circ \partial = 0$$

Donc $B_*(K) \subset Z_*(K)$ où $B_*(K) = \text{Im } \partial$ et $Z_*(K) = \ker \partial$.

$$H_*(K) = Z_*(K)/B_*(K)$$

Morphisme induit

$f : K \rightarrow L$ compatible avec les bords :
 $\partial \circ f_{\#} = f_{\#} \circ \partial$ où $f_{\#} : C_*(K) \rightarrow C_*(L)$.

On en déduit $f_{\#}(B_*(K)) \subset B_*(L)$, puis

$$f_* : H_*(K) \rightarrow H_*(L)$$

Exemple : $K \subset L$ donne $H_*(K) \rightarrow H_*(L)$.

Morphisme induit

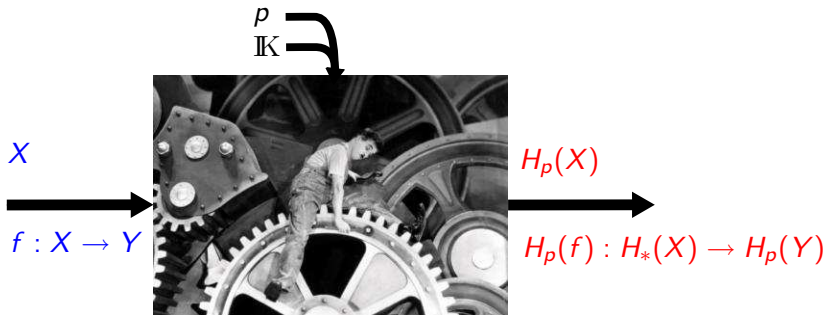
$f : K \rightarrow L$ compatible avec les bords :
 $\partial \circ f_{\#} = f_{\#} \circ \partial$ où $f_{\#} : C_*(K) \rightarrow C_*(L)$.

On en déduit $f_{\#}(B_*(K)) \subset B_*(L)$, puis

$$f_* : H_*(K) \rightarrow H_*(L)$$

Exemple : $K \subset L$ donne $H_*(K) \rightarrow H_*(L)$.

Le foncteur homologique



$$H_p(\text{Id}) = \text{Id} \text{ et } H_p(f \circ g) = H_p(f) \circ H_p(g)$$

H_p facile à calculer si X et f sont simpliciaux.

Exemple

$$K_1 \subset K_2 \subset \dots \subset K_n = K$$

$$\downarrow H_*$$

$$H_*(K_1) \rightarrow H_*(K_2) \rightarrow \dots \rightarrow H_*(K_n)$$

Filtration $\xrightarrow{H_*}$ chaîne d'applications linéaires.

On remplace l'étude de la filtration par celle des chaîne d'applications linéaires.

Exemple

$$K_1 \subset K_2 \subset \dots \subset K_n = K$$

$$\Downarrow H_*$$

$$H_*(K_1) \rightarrow H_*(K_2) \rightarrow \dots \rightarrow H_*(K_n)$$

Filtration $\xrightarrow{H_*}$ chaîne d'applications linéaires.

On remplace l'étude de la filtration par celle des chaîne d'applications linéaires.

Exemple

$$K_1 \subset K_2 \subset \dots \subset K_n = K$$

$$\Downarrow H_*$$

$$H_*(K_1) \rightarrow H_*(K_2) \rightarrow \dots \rightarrow H_*(K_n)$$

Filtration $\xrightarrow{H_*}$ chaîne d'applications linéaires.

On remplace l'étude de la filtration par celle des chaîne d'applications linéaires.

Morphisme et somme directe de chaînes

$$(f_i) : E_1 \xrightarrow{f_1} E_2 \xrightarrow{f_2} \dots \xrightarrow{f_{n-1}} E_n$$

$$(g_i) : F_1 \xrightarrow{g_1} F_2 \xrightarrow{g_2} \dots \xrightarrow{g_{n-1}} F_n$$

$(f_i) \simeq (g_i)$ si les ϕ_i sont des isomorphismes.

$$(f_i) \oplus (g_i) : E_1 \oplus F_1 \xrightarrow{f_1 \oplus g_1} E_2 \oplus F_2 \xrightarrow{f_2 \oplus g_2} \dots \xrightarrow{f_{n-1} \oplus g_{n-1}} E_n \oplus F_n$$

Morphisme et somme directe de chaînes

$$\begin{array}{ccccccc}
 (f_i) : & E_1 & \xrightarrow{f_1} & E_2 & \xrightarrow{f_2} & \cdots & \xrightarrow{f_{n-1}} & E_n \\
 & \downarrow \phi_1 & & \downarrow \phi_2 & & & & \downarrow \phi_n \\
 (g_i) : & F_1 & \xrightarrow{g_1} & F_2 & \xrightarrow{g_2} & \cdots & \xrightarrow{g_{n-1}} & F_n
 \end{array}$$

$(f_i) \simeq (g_i)$ si les ϕ_i sont des isomorphismes.

$$(f_i) \oplus (g_i) : E_1 \oplus F_1 \xrightarrow{f_1 \oplus g_1} E_2 \oplus F_2 \xrightarrow{f_2 \oplus g_2} \cdots \xrightarrow{f_{n-1} \oplus g_{n-1}} E_n \oplus F_n$$

Morphisme et somme directe de chaînes

$$\begin{array}{ccccccc}
 (f_i) : & E_1 & \xrightarrow{f_1} & E_2 & \xrightarrow{f_2} & \cdots & \xrightarrow{f_{n-1}} & E_n \\
 & \downarrow \phi_1 & & \downarrow \phi_2 & & & & \downarrow \phi_n \\
 (g_i) : & F_1 & \xrightarrow{g_1} & F_2 & \xrightarrow{g_2} & \cdots & \xrightarrow{g_{n-1}} & F_n
 \end{array}$$

$(f_i) \simeq (g_i)$ si les ϕ_i sont des isomorphismes.

$$(f_i) \oplus (g_i) : E_1 \oplus F_1 \xrightarrow{f_1 \oplus g_1} E_2 \oplus F_2 \xrightarrow{f_2 \oplus g_2} \cdots \xrightarrow{f_{n-1} \oplus g_{n-1}} E_n \oplus F_n$$

Morphisme et somme directe de chaînes

$$\begin{array}{ccccccc}
 (f_i) : & E_1 & \xrightarrow{f_1} & E_2 & \xrightarrow{f_2} & \dots & \xrightarrow{f_{n-1}} & E_n \\
 & \downarrow \phi_1 & & \downarrow \phi_2 & & & & \downarrow \phi_n \\
 (g_i) : & F_1 & \xrightarrow{g_1} & F_2 & \xrightarrow{g_2} & \dots & \xrightarrow{g_{n-1}} & F_n
 \end{array}$$

$(f_i) \simeq (g_i)$ si les ϕ_i sont des isomorphismes.

$$(f_i) \oplus (g_i) : E_1 \oplus F_1 \xrightarrow{f_1 \oplus g_1} E_2 \oplus F_2 \xrightarrow{f_2 \oplus g_2} \dots \xrightarrow{f_{n-1} \oplus g_{n-1}} E_n \oplus F_n$$

Classification des chaînes d'applications linéaires

Chaînes indécomposables :

$$\mathrm{Id}_{\mathbb{K}}[a, b] : \overset{1}{0} \longrightarrow \cdots \longrightarrow 0 \longrightarrow \overset{a}{\mathbb{K}} \xrightarrow{\mathrm{Id}} \cdots \xrightarrow{\mathrm{Id}} \overset{b}{\mathbb{K}} \longrightarrow 0 \cdots \longrightarrow \overset{n+1}{0}$$

Décomposition canonique

$\exists!$ multi-ensemble d'intervalles, I , tq

$$(f_i)_{i \in [1, n]} \simeq \bigoplus_{[a, b] \in I} \mathrm{Id}_{\mathbb{K}}[a, b]$$

Corollaire

I est un invariant complet pour (f_i) . Ces éléments sont les intervalles de persistance de (f_i) .

Classification des chaînes d'applications linéaires

Chaînes indécomposables :

$$\mathrm{Id}_{\mathbb{K}}[a, b] : 0 \xrightarrow{1} \cdots \rightarrow 0 \xrightarrow{a} \mathbb{K} \xrightarrow{\mathrm{Id}} \cdots \xrightarrow{\mathrm{Id}} \mathbb{K} \xrightarrow{b} 0 \cdots \rightarrow 0 \xrightarrow{n+1} 0$$

Décomposition canonique

$\exists!$ multi-ensemble d'intervalles, I , tq

$$(f_i)_{i \in [1, n]} \simeq \bigoplus_{[a, b] \in I} \mathrm{Id}_{\mathbb{K}}[a, b]$$

Corollaire

I est un invariant complet pour (f_i) . Ces éléments sont les intervalles de persistance de (f_i) .

Classification des chaînes d'applications linéaires

Chaînes indécomposables :

$$\mathrm{Id}_{\mathbb{K}}[a, b] : 0 \xrightarrow{1} \cdots \xrightarrow{0} 0 \xrightarrow{a} \mathbb{K} \xrightarrow{\mathrm{Id}} \cdots \xrightarrow{\mathrm{Id}} \mathbb{K} \xrightarrow{b} 0 \cdots \xrightarrow{0} 0$$

Décomposition canonique

$\exists!$ multi-ensemble d'intervalles, I , tq

$$(f_i)_{i \in [1, n]} \simeq \bigoplus_{[a, b] \in I} \mathrm{Id}_{\mathbb{K}}[a, b]$$

Corollaire

I est un invariant complet pour (f_i) . Ces éléments sont les **intervalles de persistance** de (f_i) .

Preuves de la classification

- Par les rangs et les bases compatibles
- Par réduction à des indécomposables
- Par résultat classique d'algèbre commutative

$\mathbb{K}[t]$ -module gradué $\mathcal{M} = \bigoplus_i E_i$ avec $te_i = f_i(e_i)$.

Théorème de structure

Tout module gradué de type fini sur un PID gradué D est isomorphe (au sens gradué) à une unique décomposition de la forme

$$\bigoplus_{i=1}^n D(\alpha_i) \oplus \bigoplus_{j=1}^m (D/d_j D)(\alpha_j)$$

où $d_j \in D$ est homogène, $d_j \mid d_{j+1}$ et $\alpha_i, \alpha_j \in \mathbb{Z}$.

Unicité de la décomposition

Supposons $(f_k)_{k \in [1,n]} \simeq \bigoplus_{[i,j] \in I} \text{Id}_{\mathbb{K}}[i,j]$. On pose

- $f_{i,j} : E_i \rightarrow E_j$.
- $\beta_{i,j} = \text{rang}(f_{i,j})$.
- $n_{i,j} = \text{multiplicité de } [i,j] \text{ dans } I$.

Lemme

$$n_{i,j} = \beta_{i,j} - \beta_{i-1,j} - (\beta_{i,j+1} - \beta_{i-1,j+1})$$

Unicité de la décomposition

Supposons $(f_k)_{k \in [1,n]} \simeq \bigoplus_{[i,j] \in I} \text{Id}_{\mathbb{K}}[i,j]$. On pose

- $f_{i,j} : E_i \rightarrow E_j$.
- $\beta_{i,j} = \text{rang}(f_{i,j})$.
- $n_{i,j} = \text{multiplicité de } [i,j] \text{ dans } I$.

Lemme

$$n_{i,j} = \beta_{i,j} - \beta_{i-1,j} - (\beta_{i,j+1} - \beta_{i-1,j+1})$$

Unicité de la décomposition

Supposons $(f_k)_{k \in [1,n]} \simeq \bigoplus_{[i,j] \in I} \text{Id}_{\mathbb{K}}[i,j]$. On pose

- $f_{i,j} : E_i \rightarrow E_j$.
- $\beta_{i,j} = \text{rang}(f_{i,j})$.
- $n_{i,j} = \text{multiplicité de } [i,j] \text{ dans } I$.

Lemme

$$n_{i,j} = \beta_{i,j} - \beta_{i-1,j} - (\beta_{i,j+1} - \beta_{i-1,j+1})$$

preuve :

$$\beta_{i,j}((f_k)_k \oplus (g_k)_k) = \beta_{i,j}((f_k)_k) + \beta_{i,j}((g_k)_k) \text{ et}$$

$$\beta_{i,j}(\text{Id}_{\mathbb{K}}(k, l)) = \begin{cases} 1 & \text{si } [i,j] \subset [k, l] \\ 0 & \text{sinon.} \end{cases}$$

Unicité de la décomposition

Supposons $(f_k)_{k \in [1, n]} \simeq \bigoplus_{[i, j] \in I} \text{Id}_{\mathbb{K}}[i, j]$. On pose

- $f_{i,j} : E_i \rightarrow E_j$.
- $\beta_{i,j} = \text{rang}(f_{i,j})$.
- $n_{i,j} = \text{multiplicité de } [i, j] \text{ dans } I$.

Lemme

$$n_{i,j} = \beta_{i,j} - \beta_{i-1,j} - (\beta_{i,j+1} - \beta_{i-1,j+1})$$

preuve :

$$\beta_{i,j}((f_k)_k \oplus (g_k)_k) = \beta_{i,j}((f_k)_k) + \beta_{i,j}((g_k)_k) \text{ et}$$

$$\beta_{i,j}(\text{Id}_{\mathbb{K}}(k, l)) = \begin{cases} 1 & \text{si } [i, j] \subset [k, l] \\ 0 & \text{sinon.} \end{cases}$$

$$\implies \delta_{i,j} := \beta_{i,j} - \beta_{i-1,j} = \#\{[i, \ell] \subset I \mid \ell \geq j\}.$$

$$\implies n_{i,j} = \delta_{i,j} - \delta_{i,j+1}. \quad \square$$

Unicité de la décomposition

Supposons $(f_k)_{k \in [1,n]} \simeq \bigoplus_{[i,j] \in I} \text{Id}_{\mathbb{K}}[i,j]$. On pose

- $f_{i,j} : E_i \rightarrow E_j$.
- $\beta_{i,j} = \text{rang}(f_{i,j})$.
- $n_{i,j} = \text{multiplicité de } [i,j] \text{ dans } I$.

Lemme

$$n_{i,j} = \beta_{i,j} - \beta_{i-1,j} - (\beta_{i,j+1} - \beta_{i-1,j+1})$$

Corollaire

I est un invariant.

Existence de la décomposition

Soit $E_1 \xrightarrow{f_1} E_2 \xrightarrow{f_2} \dots \xrightarrow{f_n} E_{n+1}$.

Pour $x \in E_i$, on pose $x(j) = f_{i,j}(x)$.

Base compatible

famille $X \subset \bigcup_i E_i$ telle que pour $1 \leq i \leq n+1$

$$X(i) := \{x(i) \mid (x \in X) \wedge (x(i) \neq 0)\}$$

est une base de E_i .

L'intervalle de persistance de $x \in X$ est $I_x := \{i \mid x(i) \neq 0\}$.

Proposition

Toute chaîne admet une base compatible.

Preuve : La construire par récurrence sur la longueur de la chaîne. □

Existence de la décomposition

Soit $E_1 \xrightarrow{f_1} E_2 \xrightarrow{f_2} \dots \xrightarrow{f_n} E_{n+1}$.

Pour $x \in E_i$, on pose $x(j) = f_{i,j}(x)$.

Base compatible

famille $X \subset \bigcup_i E_i$ telle que pour $1 \leq i \leq n+1$

$$X(i) := \{x(i) \mid (x \in X) \wedge (x(i) \neq 0)\}$$

est une base de E_i .

L'intervalle de persistance de $x \in X$ est $I_x := \{i \mid x(i) \neq 0\}$.

Lemme

Si $(f_i)_i$ admet une X base compatible alors elles admettent une décomposition canonique avec $I = \{I_x \mid x \in X\}$

Preuve : Vérifier que $\bigoplus_{x \in X} \text{Id}_{\mathbb{K}}(I_x) \simeq (f_i)_i$ en définissant des morphismes à l'aide des bases $X(i)$. □

Existence de la décomposition

Soit $E_1 \xrightarrow{f_1} E_2 \xrightarrow{f_2} \dots \xrightarrow{f_n} E_{n+1}$.

Pour $x \in E_i$, on pose $x(j) = f_{i,j}(x)$.

Base compatible

famille $X \subset \bigcup_i E_i$ telle que pour $1 \leq i \leq n+1$

$$X(i) := \{x(i) \mid (x \in X) \wedge (x(i) \neq 0)\}$$

est une base de E_i .

L'intervalle de persistance de $x \in X$ est $I_x := \{i \mid x(i) \neq 0\}$.

Proposition

Toute chaîne admet une base compatible.

Preuve : La construire par récurrence sur la longueur de la chaîne. □

Persistence des sous-chaînes

$$\begin{array}{ccccccc}
 (f_i) : E_1 & \xrightarrow{f_1} & E_2 & \xrightarrow{f_2} & \dots & \xrightarrow{f_{i-1}} & E_i \xrightarrow{f_i} \dots \xrightarrow{f_{j-1}} E_j \dots \xrightarrow{f_n} E_{n+1} \\
 & & \downarrow = & & & & \downarrow = \\
 (f_j^\kappa) : & & E_{\kappa(1)} & \longrightarrow & E_{\kappa(2)} & \longrightarrow & \dots \longrightarrow E_{\kappa(m)}
 \end{array}$$

$$\begin{array}{ccccccc}
 \longrightarrow & E_i & \longrightarrow & & \dots & & \longrightarrow E_j \longrightarrow \\
 & & \searrow \text{red} & & & & \swarrow \text{blue} \\
 & & E_{\kappa(a)} & \longrightarrow & \dots & \longrightarrow & E_{\kappa(b)} \longrightarrow
 \end{array}$$

$\lambda(i) = a$ et $\mu(j) = b$.

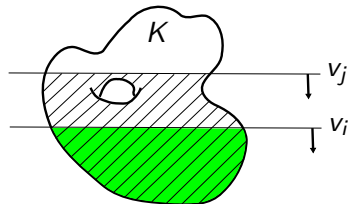
Lemme

$$I_\kappa = \{[\lambda(i), \mu(j)] \mid [i, j] \in I \text{ et } \lambda(i) \leq \mu(j)\}.$$

Preuve : Le vérifier pour les chaînes indécomposables. □

Fonctions et persistance

Soit $f : K \rightarrow \mathbb{R}$, croissante sur $K : \sigma \prec \tau \implies f(\sigma) \leq f(\tau)$.



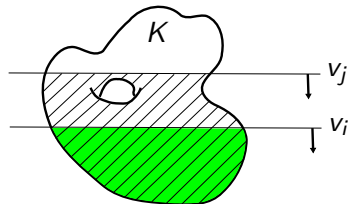
Soit $v_1 < v_2 < \dots < v_n$ les valeurs de f , $v_{n+1} = +\infty$ et $K_i = \{f \leq v_i\}$.

On a $i < j \implies K_i \subset K_j$

Fonction $f \longrightarrow$ filtration $\mathcal{K}_f \xrightarrow{H_p}$ chaîne d'applications linéaires
 \longrightarrow intervalles de persistance $I(\mathcal{K}_f)$

Fonctions et persistance

Soit $f : K \rightarrow \mathbb{R}$, croissante sur $K : \sigma \prec \tau \implies f(\sigma) \leq f(\tau)$.



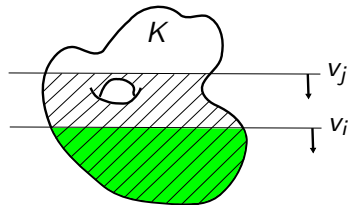
Soit $v_1 < v_2 < \dots < v_n$ les valeurs de f , $v_{n+1} = +\infty$ et $K_i = \{f \leq v_i\}$.

On a $i < j \implies K_i \subset K_j$

Fonction $f \longrightarrow$ filtration $\mathcal{K}_f \xrightarrow{H_p}$ chaîne d'applications linéaires
 \longrightarrow intervalles de persistance $I(\mathcal{K}_f)$

Fonctions et persistance

Soit $f : K \rightarrow \mathbb{R}$, croissante sur $K : \sigma \prec \tau \implies f(\sigma) \leq f(\tau)$.



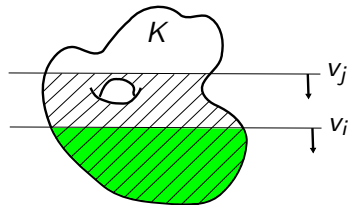
Soit $v_1 < v_2 < \dots < v_n$ les valeurs de f , $v_{n+1} = +\infty$ et $K_i = \{f \leq v_i\}$.

On a $i < j \implies K_i \subset K_j$

Fonction $f \longrightarrow$ filtration $\mathcal{K}_f \xrightarrow{H_p}$ chaîne d'applications linéaires
 \longrightarrow intervalles de persistance $I(\mathcal{K}_f)$

Fonctions et persistance

Soit $f : K \rightarrow \mathbb{R}$, croissante sur $K : \sigma \prec \tau \implies f(\sigma) \leq f(\tau)$.



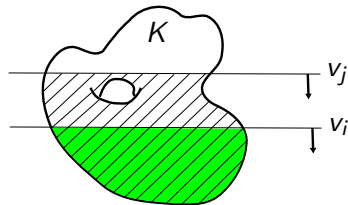
Soit $v_1 < v_2 < \dots < v_n$ les valeurs de f , $v_{n+1} = +\infty$ et $K_i = \{f \leq v_i\}$.

On a $i < j \implies K_i \subset K_j$

Fonction $f \longrightarrow$ filtration $\mathcal{K}_f \xrightarrow{H_p}$ chaîne d'applications linéaires
 \longrightarrow intervalles de persistance $I(\mathcal{K}_f)$

Fonctions et persistance

Soit $f : K \rightarrow \mathbb{R}$, croissante sur $K : \sigma \prec \tau \implies f(\sigma) \leq f(\tau)$.

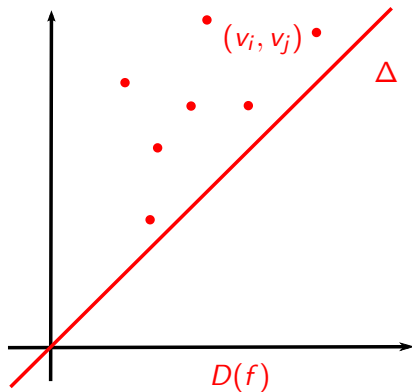


Soit $v_1 < v_2 < \dots < v_n$ les valeurs de f , $v_{n+1} = +\infty$ et $K_i = \{f \leq v_i\}$.

On a $i < j \implies K_i \subset K_j$

Fonction $f \longrightarrow$ filtration $\mathcal{K}_f \xrightarrow{H_p}$ chaîne d'applications linéaires
 \longrightarrow intervalles de persistance $I(\mathcal{K}_f)$

Diagramme de persistance



Définition

$$D(f) = \{(v_i, v_j) \mid [i, j[\in I(\mathcal{K}_f)\} \cup \Delta^\infty.$$

Stabilité du diagramme de persistance

Soient $f, g : K \rightarrow \mathbb{R}$ croissantes. On pose :

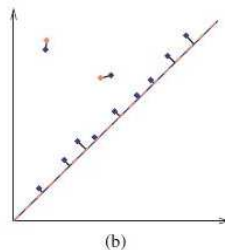
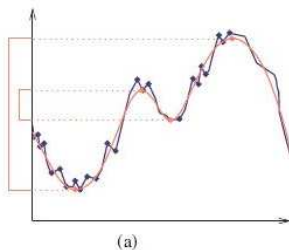
$$\|f - g\|_{\infty} = \sup_{\sigma \in K} |f(\sigma) - g(\sigma)|$$

et

$$d(D, D') = \inf_{\phi} \sup_{p \in D} \|p - \phi(p)\|_{\infty}$$

Théorème de stabilité [CEH07]

$$d(D(f), D(g)) \leq \|f - g\|_{\infty}$$



Preuve du théorème de stabilité

Définition

f **compatible** avec $\mathcal{K} = K_1 \subset K_2 \subset \dots \subset K_m$ si \mathcal{K}_f sous-filtration de \mathcal{K} . On pose alors

$$f_{\mathcal{K}}(i) = f(K_i \setminus K_{i-1}) \text{ et } f_{\mathcal{K}}(m+1) = +\infty$$

Diagramme relatif

Si f compatible avec \mathcal{K} , on pose

$$D(f, \mathcal{K}) = \Delta^\infty \cup \{(f_{\mathcal{K}}(i), f_{\mathcal{K}}(j+1))\}_{[i,j] \in I(\mathcal{K})}$$

Lemme

$$D(f) = D(f, \mathcal{K}).$$

Preuve : Appliquer le lemme des sous-chaînes et remarquer que les points $(f_{\mathcal{K}}(i), f_{\mathcal{K}}(j+1))$ avec $f_{\mathcal{K}}(i) = f_{\mathcal{K}}(j+1)$ sont absorbés par Δ^∞ . □

Preuve du théorème de stabilité

Définition

f **compatible** avec $\mathcal{K} = K_1 \subset K_2 \subset \dots \subset K_m$ si \mathcal{K}_f sous-filtration de \mathcal{K} . On pose alors

$$f_{\mathcal{K}}(i) = f(K_i \setminus K_{i-1}) \text{ et } f_{\mathcal{K}}(m+1) = +\infty$$

Diagramme relatif

Si f compatible avec \mathcal{K} , on pose

$$D(f, \mathcal{K}) = \Delta^\infty \cup \{(f_{\mathcal{K}}(i), f_{\mathcal{K}}(j+1))\}_{[i,j] \in I(\mathcal{K})}$$

Lemme

$$D(f) = D(f, \mathcal{K}).$$

Preuve : Appliquer le lemme des sous-chaînes et remarquer que les points $(f_{\mathcal{K}}(i), f_{\mathcal{K}}(j+1))$ avec $f_{\mathcal{K}}(i) = f_{\mathcal{K}}(j+1)$ sont absorbés par Δ^∞ . □

Preuve du théorème de stabilité (cont')

on pose $f_t = f + t(g - f)$.

$\forall \sigma, \tau \in K, \exists u$ tel que : $\text{signe}(f_t(\sigma) - f_t(\tau))$ est constant pour $t \in [0, u]$ (resp. pour $t \in [u, 1]$).

Donc, $\exists 0 = t_0 < t_1 < \dots < t_r = 1$ tel que l'ordre des f_t -valeurs des simplexes est constant sur chaque $[t_i, t_{i+1}]$. On en déduit \mathcal{K}_i compatible avec toutes les f_t pour $t \in [t_i, t_{i+1}]$. d'où

$$D(f_t) = D(f_t, \mathcal{K}_i) = \Delta^\infty \cup \{(f_t(\sigma_a), f_t(\sigma_{b+1}))\}_{[a,b] \in I(\mathcal{K}_i)}$$

et

$$d(D(f_{t_i}), D(f_{t_{i+1}})) \leq (t_{i+1} - t_i) \|f - g\|_\infty$$

et

$$d(D(f), D(g)) \leq \sum_i (t_{i+1} - t_i) \|f - g\|_\infty = \|f - g\|_\infty$$



Preuve du théorème de stabilité (cont')

on pose $f_t = f + t(g - f)$.

$\forall \sigma, \tau \in K, \exists u$ tel que : $\text{signe}(f_t(\sigma) - f_t(\tau))$ est constant pour $t \in [0, u]$ (resp. pour $t \in [u, 1]$).

Donc, $\exists 0 = t_0 < t_1 < \dots < t_r = 1$ tel que l'ordre des f_t -valeurs des simplexes est constant sur chaque $[t_i, t_{i+1}]$. On en déduit \mathcal{K}_i compatible avec toutes les f_t pour $t \in [t_i, t_{i+1}]$. d'où

$$D(f_t) = D(f_t, \mathcal{K}_i) = \Delta^\infty \cup \{(f_t(\sigma_a), f_t(\sigma_{b+1}))\}_{[a,b] \in I(\mathcal{K}_i)}$$

et

$$d(D(f_{t_i}), D(f_{t_{i+1}})) \leq (t_{i+1} - t_i) \|f - g\|_\infty$$

et

$$d(D(f), D(g)) \leq \sum_i (t_{i+1} - t_i) \|f - g\|_\infty = \|f - g\|_\infty$$



Preuve du théorème de stabilité (cont')

on pose $f_t = f + t(g - f)$.

$\forall \sigma, \tau \in K, \exists u$ tel que : $\text{signe}(f_t(\sigma) - f_t(\tau))$ est constant pour $t \in [0, u]$ (resp. pour $t \in [u, 1]$).

Donc, $\exists 0 = t_0 < t_1 < \dots < t_r = 1$ tel que l'ordre des f_t -valeurs des simplexes est constant sur chaque $[t_i, t_{i+1}]$. On en déduit \mathcal{K}_i compatible avec toutes les f_t pour $t \in [t_i, t_{i+1}]$. d'où

$$D(f_t) = D(f_t, \mathcal{K}_i) = \Delta^\infty \cup \{(f_t(\sigma_a), f_t(\sigma_{b+1}))\}_{[a,b] \in I(\mathcal{K}_i)}$$

et

$$d(D(f_{t_i}), D(f_{t_{i+1}})) \leq (t_{i+1} - t_i) \|f - g\|_\infty$$

et

$$d(D(f), D(g)) \leq \sum_i (t_{i+1} - t_i) \|f - g\|_\infty = \|f - g\|_\infty$$



Preuve du théorème de stabilité (cont')

on pose $f_t = f + t(g - f)$.

$\forall \sigma, \tau \in K, \exists u$ tel que : $\text{signe}(f_t(\sigma) - f_t(\tau))$ est constant pour $t \in [0, u]$ (resp. pour $t \in [u, 1]$).

Donc, $\exists 0 = t_0 < t_1 < \dots < t_r = 1$ tel que l'ordre des f_t -valeurs des simplexes est constant sur chaque $[t_i, t_{i+1}]$. On en déduit \mathcal{K}_i compatible avec toutes les f_t pour $t \in [t_i, t_{i+1}]$. d'où

$$D(f_t) = D(f_t, \mathcal{K}_i) = \Delta^\infty \cup \{(f_t(\sigma_a), f_t(\sigma_{b+1}))\}_{[a,b] \in I(\mathcal{K}_i)}$$

et

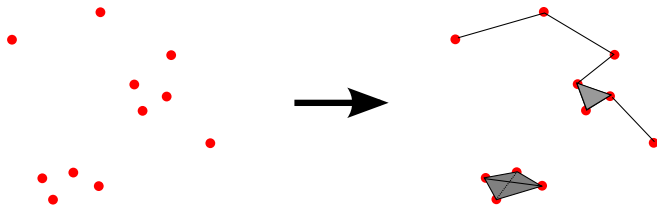
$$d(D(f_{t_i}), D(f_{t_{i+1}})) \leq (t_{i+1} - t_i) \|f - g\|_\infty$$

et

$$d(D(f), D(g)) \leq \sum_i (t_{i+1} - t_i) \|f - g\|_\infty = \|f - g\|_\infty$$



Inférence topologique



On connaît P un échantillon de points sur une variété $\mathcal{M} \subset \mathbb{R}^n$.
On cherche à évaluer la topologie de \mathcal{M} à partir de P .

On construit une famille croissante de complexes sur P : Complexe de Rips, alpha-complexe, complexe de Čech ... On sélectionne les intervalles de persistance “longs”.

Inférence topologique

- cercle
- sphère
- tore

Animations extraites de : Singh, G., Memoli, F., Ishkhanov, T., Sapiro, G., Carlsson, G., and Ringach, D. L. **Topological analysis of population activity in visual cortex.** *Journal of Vision*, 8(8), 2008.

Filtrations simples

Soit $\mathcal{K} : K_1 \subset K_2 \subset \dots \subset K_n$ où $K_i = \mathbb{K}_{i-1} \cup \sigma_i$.

Formule du rang :

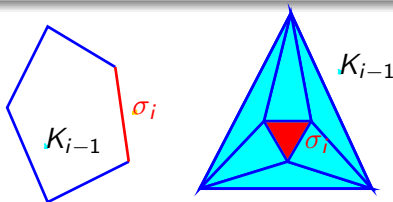
$$\dim C_*(K_i) = \dim Z_*(K_i) + \dim B_*(K_i)$$

$$\dim C_*(K_i) = \dim C_*(K_{i-1}) + 1 \implies$$

Soit $\dim Z_*(K_i)$ soit $\dim B_*(K_i)$ augmente de 1.

Définition

σ_i (resp. i) est dit **positif** si $\dim Z_*(K_i) = \dim Z_*(K_{i-1}) + 1$ et **négatif** sinon. On note $\mathcal{P}(\mathcal{K})$ et $\mathcal{N}(\mathcal{K})$ les indices positifs resp. négatifs.



Structure des intervalles

Lemme

Tout intervalle de $I(\mathcal{K})$ est de la forme $[i, j]$ avec
 $(i, j + 1) \in \mathcal{P}(\mathcal{K}) \times (\mathcal{N}(\mathcal{K}) \cup \{n + 1\})$.

- Un indice positif est la borne inférieure d'un unique intervalle de $I(\mathcal{K})$,
- un indice négatif j est tel que $j - 1$ est la borne supérieure d'un unique intervalle de $I(\mathcal{K})$.

Preuve : Utiliser $\dim H_*(K_i) =$ nombre d'intervalles de $I(\mathcal{K})$ contenant i . □

Définition

On note $N : \mathcal{N}(\mathcal{K}) \rightarrow \mathcal{P}(\mathcal{K}), j \mapsto N(j)$ tel que $[N(j), j - 1] \in I(\mathcal{K})$

Base compatible des bords

Définition

Base compatible des bords est une famille

$\mathcal{B}(\mathcal{K}) = \{x_j\}_{j \in J} \subset Z_*(K)$, avec $J \subset [1, n]$, telle que :

- ① $\forall i \in [1, n], \{x_j \mid (j \in J) \wedge (j \leq i)\}$ est une base de $B_*(K_i)$,
- ② L'application $\nu : J \rightarrow [1, n], j \mapsto$ l'indice max des simplexes de x_j est *injective*.

Lemme

$\nu = N$

Preuve : Soit $X = \{[z_i]_i\}_{i \in \mathcal{P}(\mathcal{K})}$ avec

- $z_i = x_j$ si $j = \nu(j)$,
- choisir z_i tel que $Z(K_i) = Z(K_{i-1}) \oplus \mathbb{K}z_i$ sinon.

X est une base compatible avec les bons intervalles. □

Base compatible des bords

Définition

Base compatible des bords est une famille

$\mathcal{B}(\mathcal{K}) = \{x_j\}_{j \in J} \subset Z_*(K)$, avec $J \subset [1, n]$, telle que :

- ① $\forall i \in [1, n], \{x_j \mid (j \in J) \wedge (j \leq i)\}$ est une base de $B_*(K_i)$,
- ② L'application $\nu : J \rightarrow [1, n], j \mapsto$ l'indice max des simplexes de x_j est *injective*.

Lemme

$\nu = N$

Preuve : Soit $X = \{[z_i]_i\}_{i \in \mathcal{P}(\mathcal{K})}$ avec

- $z_i = x_j$ si $j = \nu(j)$,
- choisir z_i tel que $Z(K_i) = Z(K_{i-1}) \oplus \mathbb{K}z_i$ sinon.

X est une base compatible avec les bons intervalles. □

Algorithme

Soit $\mathcal{K}' : K_1 \subset K_2 \subset \dots \subset K_{n-1}$ et $\mathcal{B}(\mathcal{K}')$ une base compatible.

Écrire

$$\partial\sigma_n = \sum_{x \in \mathcal{B}(\mathcal{K}')} \alpha_x x + y$$

avec

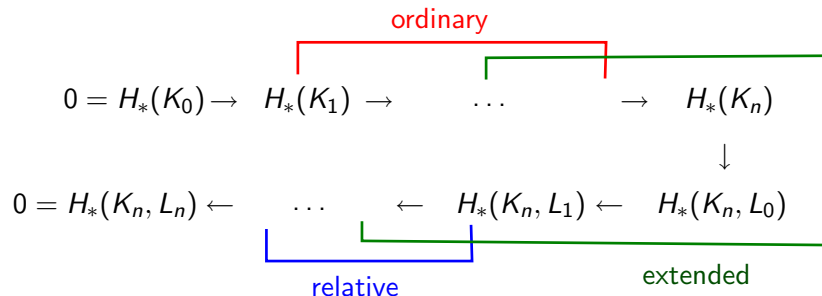
- ou bien $y = 0$,
- ou bien $\nu(y) \notin \nu(\mathcal{B}(\mathcal{K}'))$.

Alors $y = 0 \implies B_*(K_n) = B_*(K_{n-1})$ donc $\mathcal{B}(\mathcal{K}')$ base compatible pour \mathcal{K} , sinon c'est vrai pour $\mathcal{B}(\mathcal{K}') \cup \{y\}$

Puisque $\mathcal{B}(\mathcal{K}')$ échelonnée, on peut appliquer le pivot de Gauss.
Complexité en $O(n^3)$.

Diagramme étendu [CEH08]

On considère la chaîne induite par inclusion des paires



Si la filtration provient d'une fonction f on obtient
 $Dgm(f)$, $Ord(f)$, $Ext(f)$, $Rel(f)$

Propriétés

Stabilité [CEH08]

$$d(Dgm(f), Dgm(g)) \leq \|f - g\|_{\infty}$$

Symétrie [CdSM09]

$$Ord_p(f) = Rel_{p+1}^0(-f)$$

$$Ext_p(f) = Ext_p^R(-f)$$

$$Rel_p(f) = Rel_{p+1}^0(-f)$$

Dualité [CEH08]

Si f est définie sur une d -variété

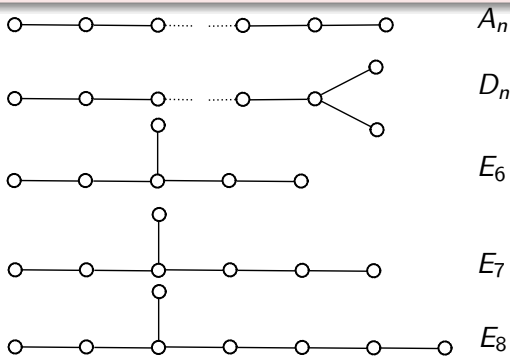
$$Dgm_p(f) = Dgm_{d-p}^T(f)$$

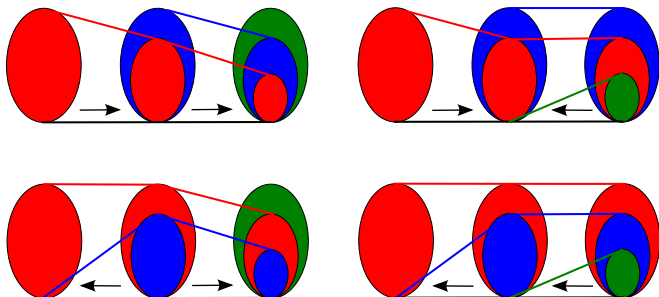
Représentations des quivers

Un quiver est de *type fini* s'il admet un nombre fini de représentations indécomposables,

Théorème de Gabriel

Les quivers de type fini sont ceux de la liste ci-dessous



Cas de A_n [CdS08]

Pour un type donné les indécomposables sont :

$$\mathrm{Id}_{\mathbb{K}}[a, b] : 0 \overset{1}{\longleftrightarrow} \dots \longleftrightarrow 0 \overset{a}{\longleftrightarrow} \mathbb{K} \overset{\mathrm{Id}}{\longleftrightarrow} \dots \overset{b}{\longleftrightarrow} \mathbb{K} \longleftrightarrow 0 \dots \overset{n}{\longleftrightarrow} 0$$

Soit \mathcal{R}_i la filtration droite de la suite tronquée en i et $(b_1^i, b_2^i, \dots, b_i^i)$ les *dates de naissance* associées.

Exemple : $B(\bullet \leftarrow \bullet \rightarrow \bullet) = (2, 1, 3)$

Décomposition canonique

Si $\mathcal{R}_i = (R_0, R_1, \dots, R_i)$ on pose
 $\dim \mathcal{R}_i = (\dim R_0, \dim R_1/R_0, \dots, \dim R_i/R_{i-1}).$

On écrit

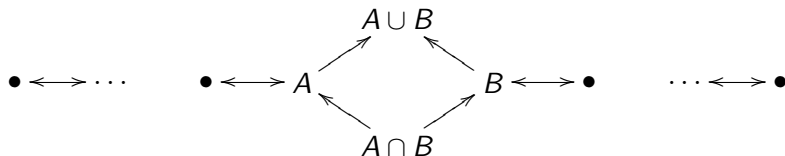
$$(c_0^i, c_1^i, \dots, c_i^i) = \begin{cases} \dim(\mathcal{R}_i \cap \ker f_i) & \text{si } f_i = \rightarrow \\ \dim \mathcal{R}_i - \dim(\mathcal{R}_i \cap \operatorname{Im} f_i) & \text{si } f_i = \leftarrow \end{cases}$$

Décomposition canonique [CdS08]

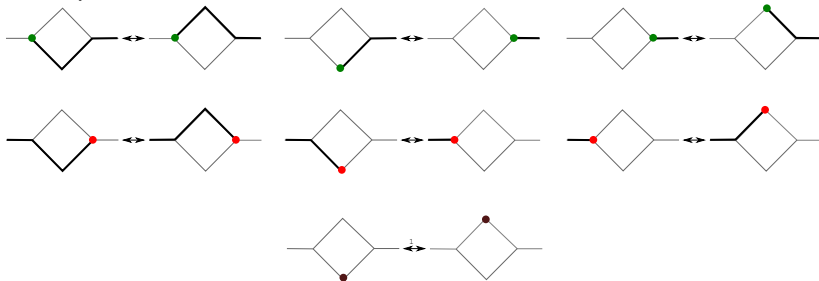
Une chaîne de type fixée est isomorphe à

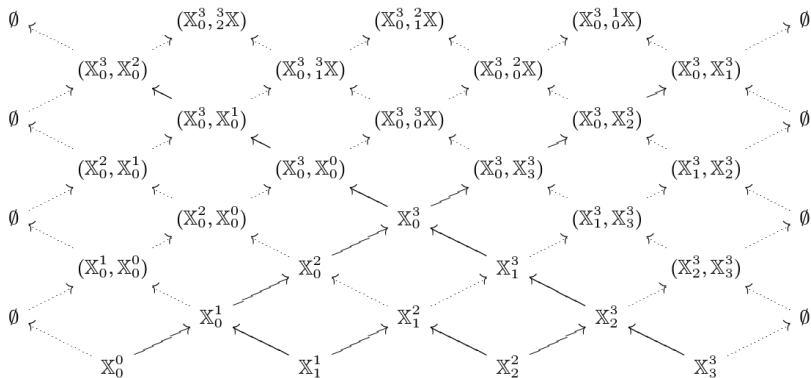
$$\bigoplus_{1 \leq i \leq k \leq n} c_i^k \operatorname{Id}_{\mathbb{K}}(b_i^k, k)$$

Principe du carreau [CdS08]



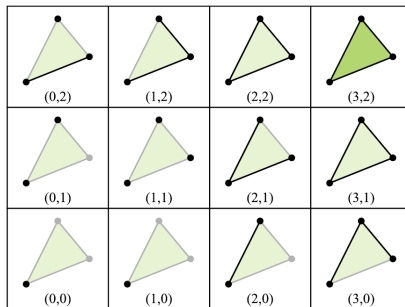
Correspondance entre intervalles :



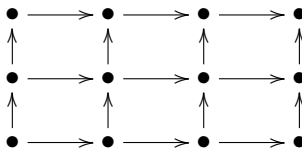


Toutes les chaines gauche-droite comportent la même information.

Multifiltration [CS09]



(fig. extraite de Carlsson et Zomorodian 09)



équivalent à un module 2-gradué sur $\mathbb{K}[t_1, t_2]$ finiment engendré.

Quelques notes sur la persistance :

<http://www.gipsa-lab.fr/~francis.lazarus/Enseignement/persistence-homol.pdf>