

SEMINARVORTRAG: PERSISTENZ

HELGE BÖSCHEN

ZUSAMMENFASSUNG. In diesem Vortrag wird Persistente Homologie definiert, es werden weiter Persistenz-Diagramme, Barcodes und Beispiele diskutiert. Der Vortrag wurde am 16.06.2010 gehalten im Seminar 'Algorithmische Topologie' an der Uni Münster gehalten.

1. PERSISTENZ

Definition 1.1 (Filtrierung). Sei K ein simplizialer Komplex und $n \in \mathbb{N}$. Eine Filtrierung \mathcal{K} von K ist eine Folge von Inklusionen

$$\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K.$$

Indem wir die Filtrierung verfeinern, können wir annehmen, dass $K_i - K_{i-1}$ immer aus genau einem Simplex besteht.

Beispiel 1.2 (Čech-Komplex). Sei $S \subset \mathbb{R}^d$ eine endliche Punktmenge und $\mathcal{C}(S, r)$ der zugehörige Čech-Komplex für einen Radius $r \in [0, \infty]$. Für $r_1 < r_2 < \dots < r_n = r$ gibt

$$\emptyset \subseteq \mathcal{C}(S, r_1) \subseteq \dots \subseteq \mathcal{C}(S, r_n)$$

Anlass zu einer Filtrierung von $\mathcal{C}(S, r)$.

Die Inklusionen einer Filtrierung geben Anlass zu Abbildungen der Homologiegruppen $f_p^{i,j} : H_p(K_i) \rightarrow H_p(K_j)$. Persistente Homologie untersucht, wie sich Homologieklassen unter diesen Abbildungen verhalten.

Definition 1.3 (Persistente Homologie). Die p -ten persistenten Homologiegruppen einer Filtrierung $\mathcal{K} = \{K_i\}_{1 \leq i \leq n}$ sind die Bilder der von den Inklusionen induzierten Abbildungen

$$H_p^{i,j} := H_p^{i,j}(\mathcal{K}) := \text{im } f_p^{i,j} \text{ für } 0 \leq i \leq j \leq n.$$

Die p -ten persistenten Bettizahlen sind definiert als $\beta_p^{i,j} = \text{rg } H_p^{i,j}$.

Bemerkung 1.4. Es gilt:

- (1) $H_p^{i,i} = H_p(K_i)$
- (2) $H_p^{i,j} = Z_p(K_i) / (B_p(K_j) \cap Z_p(K_i))$

Wir sagen, $\gamma \in H_p(K_i)$ ist in K_i geboren, falls $\gamma \notin H_p^{i-1,i}$ und γ stirbt in K_j , falls γ in K_j mit einer älteren Homologieklass verschmilzt, also wenn $f_p^{i,j-1}(\gamma) \notin H_p^{i-1,j-1}$ aber $f_p^{i,j}(\gamma) \in H_p^{i-1,j}$ (siehe Figur 1).

Beispiel 1.5. Betrachte Abbildung 2. Der 0-Zykel aus K_1 stirbt nie, andere 0-Zykel sterben spätestens in K_9 . Der 1-Zykel aus K_{10} stirbt in K_{14} . Würde in K_{14} der andere 2-Simplex eingefügt werden, würden sich die persistenten, nicht aber die gewöhnlichen Bettizahlen verändern.

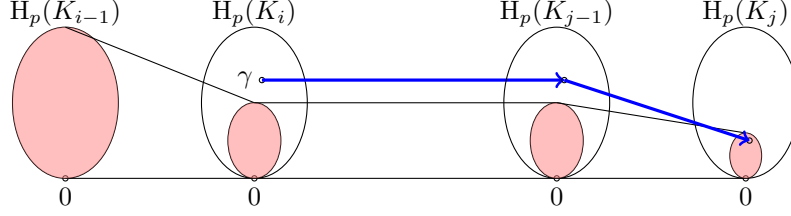


ABBILDUNG 1. Eine persistente Homologieklasse

Bemerkung 1.6 (Algebraische Struktur). Ein Persistenzkomplex ist eine Sequenz von R -Kettenkomplexen $C = (C_*^i)$ mit Kettenabbildungen $x : C_*^i \rightarrow C_*^{i+1}$. Die persistente Homologie $H_*^{i,j}(C)$ wird als Bild des induzierten Homomorphismus $x_* : H_*(C_*^i) \rightarrow H_*(C_*^j)$ definiert. Es trägt C eine $R[x]$ -Struktur, die sich auf $H_*(C)$ überträgt. Betrachten wir einen endlichen Persistenzkomplex C und Koeffizienten in einem Körper F , können wir, da $F[x]$ ein Hauptidealring ist, den Struktursatz für endlich erzeugte Moduln anwenden:

$$H_*(C; F) \cong \bigoplus_i x^{t_i} \cdot F[x] \oplus \left(\bigoplus_j x^{r_j} \cdot (F[x]/(x^{s_j} \cdot F[x])) \right).$$

Die freien Anteile entsprechen den Homologieklassen die zum Zeitpunkt t_i geboren werden und nicht sterben, die Torsionsanteile entsprechen den Klassen, die zum Zeitpunkt r_j geboren werden und zum Zeitpunkt $r_j + s_j$ sterben.

2. BARCODES

Das persistente Analogon zu Bettizahlen sind Barcodes. Ein Barcode ist eine Sammlung von horizontalen Linien in der Ebene, deren horizontale Achse zur Parametrisierung korrespondiert und deren vertikale Achse eine beliebige Ordnung der Erzeuger der Homologieklassen repräsentiert. Jede Linie repräsentiert die Lebensdauer einer Homologieklasse, siehe Abbildung 4. Die persistenten Bettizahlen lassen sich wie folgt daraus ablesen:

Satz 2.1 ([5]). Die persistente Bettizahl $\beta_p^{i,j}$ ist gleich der Anzahl der Linien im Barcode, die das Intervall $[i, j]$ überdecken.

Barcodes sind - wie Bettizahlen - ein sehr grobes topologisches Maß, ihre Stärke liegt darin, dass sie gegenüber Rauschen sehr resistent sind.

3. PERSISTENZ-DIAGRAMME

Für die algorithmische Berechnung persistenter Bettizahlen ist eine andere Darstellung in der reellen Ebene \mathbb{R}^2 nützlich. Wir definieren $\mu_p^{i,j}$ als die Anzahl der unabhängigen p -dimensionalen Homologieklassen, die in K_i geboren werden und in K_j sterben. Es ist

$$\mu_p^{i,j} = (\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j})$$

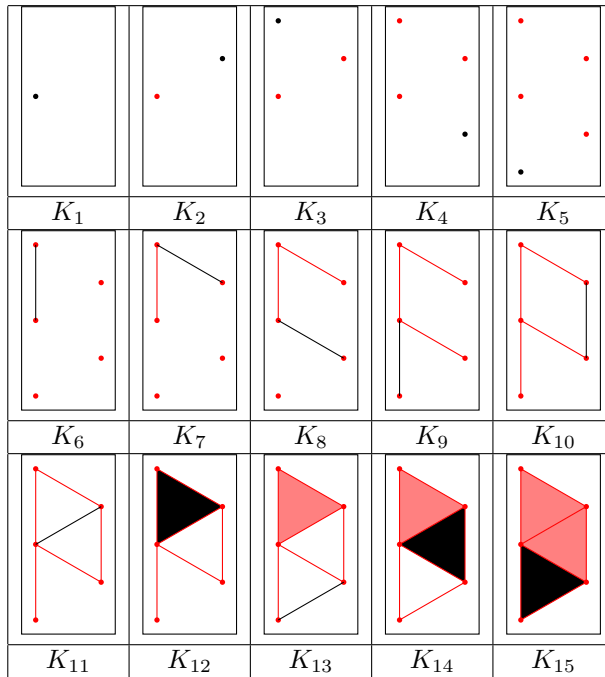


ABBILDUNG 2. Beispiel einer Filtrierung

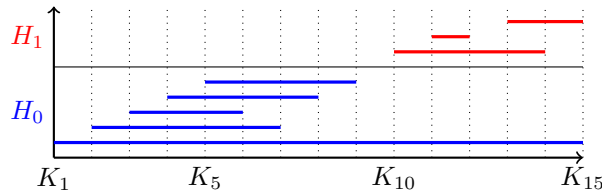


ABBILDUNG 3. Barcodes Filtrierung aus Abbildung 2

für $i < j$. Die erste Differenz zählt die Klassen, die in oder vor K_i geboren werden und in K_j sterben, die zweite Differenz zählt die Klassen, die in oder vor K_{i-1} geboren werden und in K_j sterben. Da in jedem $K_i - K_{i-1}$ nur ein Simplex ist, gilt $\mu_p^{i,j} \in \{0, 1\}$. Wir zeichnen jeden Punkt (i, j) mit $\mu_p^{i,j} = 1$ als Punkt in der Ebene und erhalten das p -te Persistenz Diagramm $\text{Dgm}_p(\mathcal{K})$.

In einer Filtrierung $\mathcal{K} = \{K_i\}_{1 \leq i \leq n}$ setzen wir $H_p(K_{n+1}) = 0$, dann stirbt jede Klasse per Definition in K_{n+1} und das Diagramm bleibt endlich.

Definition 3.1 (Persistenzdiagramm). Das p -te Persistenzdiagramm $\text{Dgm}_p(\mathcal{K})$ einer Filtrierung \mathcal{K} ist die Menge der Punkte $(i, j) \in \mathbb{R}^2$ mit $\mu_p^{i,j} = 1$.

Offensichtlich liegen alle Punkte oberhalb der Diagonalen. Aus technischen Gründen werden die Punkte auf der Diagonalen ebenfalls ins Diagramm eingezeichnet.

Die Anzahl der Punkte im linken oberen Quadranten mit der Ecke (k, l) , wobei die rechte Seite des Quadranten abgeschlossen und die untere Seite offen ist, ist die

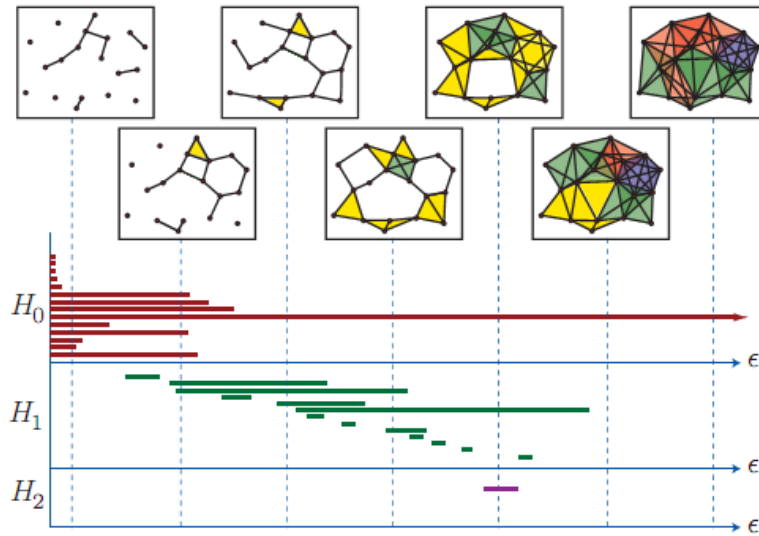


ABBILDUNG 4. Barcodes

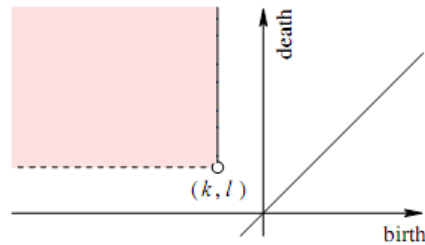


ABBILDUNG 5. Persistenzdiagramm

persistente Bettizahl $\beta_p^{k,l}$. Für $\beta_p^{k,l}$ werden nämlich genau die Klassen gezählt, die in K_i geboren werden und in K_j sterben, wobei $i \leq k$ und $j > l$ ist. Es gilt also

Lemma 3.2 (Fundamentales Lemma der Persistenten Homologie). *Für eine Filtrierung \mathcal{K} und Indizes $k \leq l$ gilt*

$$\beta_p^{k,l} = \sum_{i \leq k} \sum_{j > l} \mu_p^{i,j}.$$

Persistenzdiagramme enthalten also alle Informationen über persistente Bettizahlen, der Matrix-Reduktions-Algorithmus liefert ein solches Diagramm und damit auch die persistenten Bettizahlen.

4. SENSOR-NETZWERKE

Es sei ein Gebiet $D \subset \mathbb{R}^d$ und eine Menge von Sensoren $X \subset D$ gegeben. Jeder Sensor hat einen Überwachungsradius r_c und einen Senderadius r_s , in welchem er andere Sensoren erkennt.

Frage 4.1. Wann gilt

$$D \subset \bigcup_{x \in X} B_{r_c}(x) =: U?$$

Homologie erkennt Löcher, die Lemmata

Lemma 4.2 ([3], S. 458f). *Es gilt:*

$$\bigcup_{x \in X} B(x, r) \simeq \mathcal{C}(S, 2r) := N(\{B_r(x) : x \in X\}).$$

und

Lemma 4.3 ([1], S. 62). *Es gibt eine Kette von Inklusionen*

$$\mathcal{R}(S, r) \subseteq \mathcal{C}(S, r\sqrt{2}) \subseteq \mathcal{R}(S, r\sqrt{2}).$$

geben eine Motivation, warum persistente Homologie hier hilfreich sein wird. Der Ripskomplex ist der maximale Simpliciale Komplex, der als 1-Gerüst das 1-Gerüst des Čechkomplexes hat.

Unser Gebiet und unsere Sensoren haben folgende Eigenschaften:

- A1 Jeder Sensor hat einen Überwachungsradius r_c .
- A2 Jeder Knoten sendet eine eindeutige ID, jeder Sensor kann andere Sensoren in einem Radius r_s über ein starkes und in einem Radius r_w über ein schwaches Signal erkennen.
- A3 Die Radien erfüllen $r_c \geq r_s/\sqrt{2}$ und $r_w \geq r_s\sqrt{10}$.
- A4 Die Sensoren liegen in einer kompakten Menge $D \subset \mathbb{R}^d$. Sensoren erkennen, wenn sie zu ∂D einen Abstand $\leq r_f$ haben.
- A5 Die eingeschränkte Region $D - N_{\hat{r}}(\partial D)$ ist zusammenhängend, dabei ist

$$N_{\hat{r}}(\partial D) = \{x \in D : \|x - \partial D\| \leq \hat{r}\} \text{ mit } \hat{r} = r_f + r_s/\sqrt{2}.$$

- A6 Die Region $\Sigma = \{x \in D : \|x - \partial D\| = r_f\}$ hat internen Injektivitätsradius mindestens $r_s/\sqrt{2}$ und externen Injektivitätsradius mindestens r_s .

Die Bedingungen A1-A4 bestimmen die Kommunikationsfähigkeiten der Sensoren, A5 stellt sicher, dass die Region D keine zu 'engen' Stellen hat, A6 verhindert gewisse pathologische Fälle.

4.1. Ergebnis. Wegen A1-A4 erhalten wir das Diagramm von Kommunikationsgraphen

$$\begin{array}{ccc} G_s & \xrightarrow{\subset} & G_w \\ \uparrow \cup & & \uparrow \cup \\ F_s & \xrightarrow{\subset} & F_w \end{array}$$

Die Graphen G_s und G_w sind die Kommunikationsgraphen bezüglich der starken beziehungsweise schwachen Signale aller Sensoren, F_s und F_w die entsprechenden

Graphen der Sensoren mit einem Randabstand $\leq r_f$. Indem wir jeweils den maximalen Simplicialen Komplex bilden, der den Graph als 1-Gerüst enthält, erhalten wir die Ripskomplexe

$$\begin{array}{ccc} \mathcal{R}_s & \xrightarrow{\subset} & \mathcal{R}_w \\ \cup \uparrow & & \uparrow \cup \\ \mathcal{F}_s & \xrightarrow{\subset} & \mathcal{F}_w \end{array}$$

Satz 4.4. *Es wird $D - N_{\hat{r}}(\partial D)$ von $U = \bigcup_{x \in X} B_{r_c}(x)$ überdeckt, falls der kanonische Homomorphismus*

$$\iota_* : H_d(\mathcal{R}_s, \mathcal{F}_s) \rightarrow H_d(\mathcal{R}_w, \mathcal{F}_w)$$

nicht-trivial ist.

Beweisskizze. Eine Abschätzung liefert aufgrund von $r_c \geq r_s/\sqrt{2}$:

Lemma 4.5. *Die konvexe Hülle einer Menge von Sensoren, die einen Simplex in \mathcal{R}_s bilden, liegt vollständig in U .*

Wir definieren den erweiterten Kragen

$$\overline{N_{\hat{r}}(\partial D)} = \mathbb{R}^d - (D - N_{\hat{r}}(\partial D)).$$

Es gilt

Lemma 4.6. *Die konvexe Hülle einer Menge von Sensoren, die einen Simplex in \mathcal{F}_s bilden, liegt vollständig in $\overline{N_{\hat{r}}(\partial D)}$.*

Für den eigentlichen Beweis betrachten wir das Diagramm

$$\begin{array}{ccc} H_d(\mathcal{R}_s, \mathcal{F}_s) & \xrightarrow{\partial} & H_{d-1}(\mathcal{F}_s) \\ \downarrow \sigma_* & & \downarrow \sigma_* \\ H_d(\mathbb{R}^d, \overline{N_{\hat{r}}(\partial D)}) & \xrightarrow{\partial} & H_{d-1}(\overline{N_{\hat{r}}(\partial D)}) \end{array}$$

wobei $\sigma : \mathcal{R}_s \rightarrow D$ die Realisierung ist, bei der Ecken auf die zugehörigen Punkte und Simplexe auf die konvexen Hüllen abgebildet werden.

Sei nun $[\alpha] \in H_d(\mathcal{R}_s, \mathcal{F}_s)$ mit $\iota_*[\alpha] \neq 0$. Dann gilt

- (1) $\iota_*[\alpha] \neq 0 \Rightarrow \sigma_*\partial[\alpha] \neq 0$
- (2) $\sigma_*\partial[\alpha] \neq 0 \Rightarrow$ Es überdeckt U das Gebiet $D - N_{\hat{r}}(\partial D)$

Gilt nämlich $\sigma_*\partial[\alpha] \neq 0$, so gilt auch $\sigma_*[\alpha] \neq 0$. Wenn aber $D - N_{\hat{r}}(\partial D)$ nicht von U überdeckt wird, gibt es ein $p \in D - (N_{\hat{r}}(\partial D) \cup U)$ und da $\sigma(\mathcal{R}_s) \subset U$ gilt faktorisiert $(\mathcal{R}_s, \mathcal{F}_s) \rightarrow (\mathbb{R}^d, \overline{N_{\hat{r}}(\partial D)})$ über $(\mathbb{R}^d - p, \overline{N_{\hat{r}}(\partial D)})$, was wegen $H_d(\mathbb{R}^d - p, \overline{N_{\hat{r}}(\partial D)}) = 0$ nicht sein kann. Es muss U also $D - N_{\hat{r}}(\partial D)$ überdecken.

Um den Beweis abzuschließen nehmen wir nun $\sigma_*\partial[\alpha] = 0$ an und führen dies zu einem Widerspruch.

Zunächst führen wir ein weiteres Paar von Ripskomplexen $(\mathcal{R}_m, \mathcal{F}_m)$, das durch Wahl eines (dimensionsabhängigen) mittleren Radius zwischen r_s und r_w gewonnen

wird. Wir erhalten das Diagramm

$$\begin{array}{ccccc}
 H_d(\mathcal{R}_s) & \longrightarrow & H_d(\mathcal{R}_s, \mathcal{F}_s) & \xrightarrow{\partial} & H_{d-1}(\mathcal{F}_s) \\
 \downarrow \iota_* & & \downarrow \iota_* & & \downarrow \iota_* \\
 H_d(\mathcal{R}_m) & \longrightarrow & H_d(\mathcal{R}_m, \mathcal{F}_m) & \xrightarrow{\partial} & H_{d-1}(\mathcal{F}_m) \\
 \downarrow \iota_* & & \downarrow \iota_* & & \downarrow \iota_* \\
 H_d(\mathcal{R}_w) & \longrightarrow & H_d(\mathcal{R}_w, \mathcal{F}_w) & \xrightarrow{\partial} & H_{d-1}(\mathcal{F}_w)
 \end{array}$$

Aus $\sigma_*\partial[\alpha] = 0$ folgt $\iota_*\partial[\alpha] = 0 \in H_{d-1}(\mathcal{F}_m)$. Hier geht Bedingung A6 ein; die Lemmata 4.3 und 4.2 machen es hier möglich, aus dem Verschwinden von Homologieklassen in der geometrischen Realisierung auf das Verschwinden von Homologieklassen im Ripskomplex zu schließen, wenn man den Radius vergrößert.

Den Widerspruch erhalten wir nun durch eine Diagrammjagd, wenn wir beachten, dass $H_d(\mathcal{R}_m) \rightarrow H_d(\mathcal{R}_w)$ durch den Čech-Komplex \mathcal{C}_w faktorisiert, dessen d -te Homologie verschwindet, da er homotopieäquivalent zu einer Teilmenge von \mathbb{R}^d ist. \square

LITERATUR

- [1] H. Edelsbrunner and J.L. Harer, *Computational Topology*, AMS, Providence, 2010.
- [2] A.J. Zomorodian, *Topology for computing*, Cambridge University Press, Cambridge, 2005.
- [3] A. Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, 2002.
- [4] Robert Ghrist, *Barcodes: the persistent topology of data*, Bull. Amer. Math. Soc. **45** (2008), no.1, 61-75.
- [5] G. Carlsson and A.J. Zomorodian, *Computing persistent homology*, Discrete Comput. Geom. **33** (2005), 249-274.
- [6] V. de Silva and R. Ghrist, *Coverage in sensor networks via persistent homology*, Algebr. Geom. Topol. **7** (2007), 339-358.

E-mail address: helgeboeschen@wwu.de