Topological techniques for inference on periodic functions with phase variation

Wojciech Reise

Under the supervision of Frédéric Chazal and Bertrand Michel

Dec 6, 2023, Orsay

Add a picture of a periodic function with phase variation

TopAI chair
ANR-19-CHIA-0001

Data with phase variation

Signals with phase variation

A sample $S_1, \ldots, S_N$ has phase variation if

$$S_n = f(\gamma_n) + W_n, \qquad \text{for each } n \in \{1, \ldots, N\}, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \ldots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

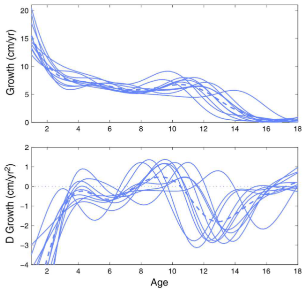


FIG. 3. *The top panel plots the growth, understood as the first derivative of height, of ten girls, and the bottom panel contains the corresponding height-acceleration or growth-derivative curves. The dashed curve in both plots is the cross-sectional mean. Both these plots indicate both phase and amplitude variability.*

Source: ???

Data with phase variation

Signals with phase variation

A sample $S_1, \dots, S_N$ has phase variation if

$$S_n = f(\gamma_n) + W_n, \qquad \text{for each } n \in \{1, \dots, N\}, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \dots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

- Computing a representative of f (Su et al. 2014)

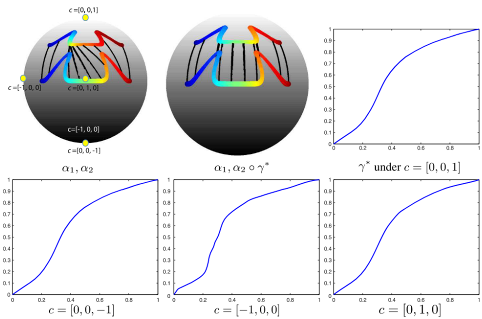


Fig. 2. *Registration of trajectories on $\mathbb{S}^2$.*

Data with phase variation

Signals with phase variation

A sample $S_1, \dots, S_N$ has phase variation if

$$S_n = f(\gamma_n) + W_n, \qquad \text{for each } n \in \{1, \dots, N\}, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \dots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

- Computing a representative of f (Su et al. 2014)
- Clustering of $S_1, \dots, S_n$ (Srivastava et al. 2011)

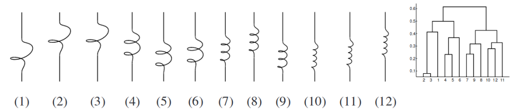


Fig. 4. A set of helices with different numbers and placements of spirals and their clustering using the elastic distance function.

Data with phase variation

Signals with phase variation

A sample $S_1, \ldots, S_N$ has phase variation if

$$S_n = f(\gamma_n) + W_n, \qquad \text{for each } n \in \{1, \ldots, N\}, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \ldots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

- Computing a representative of f (Su et al. 2014)
- Clustering of $S_1, \ldots, S_n$ (Srivastava et al. 2011)
- Synchronising curves: estimating $\gamma_n \circ \gamma_{n'}^{-1}$ (Tang and Muller 2008, Zhao and Itti 2018)

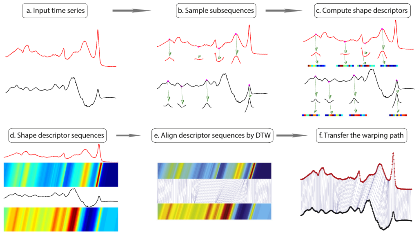


Fig. 2. Pipeline of shapeDTW. shapeDTW consists of two major steps: encode local structures by shape descriptors and align descriptor sequences by DTW. Concretely, we sample a subsequence from each temporal point, and further encode it by some shape descriptor. As a result, the original time series is converted into a descriptor sequence of the same length. Then we align two descriptor sequences by DTW and transfer the found warping path to the original time series.

Data with phase variation

Signals with phase variation

A sample $S_1, \dots, S_N$ has phase variation if

$$S_n = f(\gamma_n) + W_n, \qquad \text{for each } n \in \{1, \dots, N\}, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \dots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

- Computing a representative of f (Su et al. 2014)
- Clustering of $S_1, \dots, S_n$ (Srivastava et al. 2011)
- Synchronising curves: estimating $\gamma_n \circ \gamma_{n'}^{-1}$ (Tang and Muller 2008, Zhao and Itti 2018)

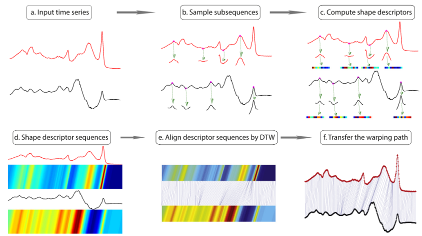


Fig. 2. Pipeline of shapeDTW. shapeDTW consists of two major steps: encode local structures by shape descriptors and align descriptor sequences by DTW. Concretely, we sample a subsequence from each temporal point, and further encode it by some shape descriptor. As a result, the original time series is converted into a descriptor sequence of the same length. Then we align two descriptor sequences by DTW and transfer the found warping path to the original time series.

Same endpoints!

For all $1 \leq n \leq N$, $\gamma_1(0) = \gamma_n(0)$ and $\gamma_1(1) = \gamma_n(1)$.

Periodic data with phase variation

Periodic function with variation

A periodic signal with phase variation is $S : [0,1] \to \mathbb{X}$

$$S = \phi(\gamma) + W_n. \tag{2}$$

where $\phi : \mathbb{R} \to \mathbb{X}$ is 1-periodic and $\gamma : [0,1] \to [0,R]$ is an increasing homeomorphism.

Phase estimation

Periodic data with phase variation

Periodic function with variation

A periodic signal with phase variation is $S : [0, 1] \to \mathbb{X}$

$$S = \phi(\gamma) + W_n. \tag{2}$$

where $\phi : \mathbb{R} \to \mathbb{X}$ is 1-periodic and $\gamma : [0, 1] \to [0, R]$ is an increasing homeomorphism.

Phase estimation

- Instantaneous phase estimation (Boashash 1992)

Periodic data with phase variation

Periodic function with variation

A periodic signal with phase variation is $S : [0,1] \to \mathbb{X}$

$$S = \phi(\gamma) + W_n. \quad (2)$$

where $\phi : \mathbb{R} \to \mathbb{X}$ is 1-periodic and $\gamma : [0,1] \to [0,R]$ is an increasing homeomorphism.

Phase estimation

- Instantaneous phase estimation (Boashash 1992)
- Zero-crossings

Periodic data with phase variation

Periodic function with variation

A periodic signal with phase variation is $S : [0, 1] \to \mathbb{X}$

$$S = \phi(\gamma) + W_n. \tag{2}$$

where $\phi : \mathbb{R} \to \mathbb{X}$ is 1-periodic and $\gamma : [0, 1] \to [0, R]$ is an increasing homeomorphism.

Phase estimation

- Instantaneous phase estimation (Boashash 1992) (Kennedy, Roth, and Scrofani 2018)
- Zero-crossings (Khasawneh and Munch 2018, Tanweer, Khasawneh, and Munch 2023)

Topological data analysis for time series data

Persistent homology from Topological data analysis (TDA) quantifies the *shape of data*.

Time series

- Detecting periodicity (Perea, Munch, and Khasawneh 2019)
- Robust zero-crossings (Khasawneh and Munch 2018)
- The number of periods in the signal is reflected by the number of prominent features (Bois et al. 2022.)

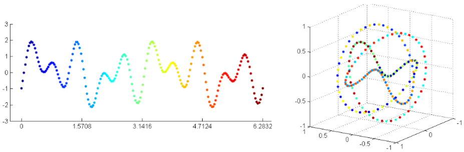


FIGURE 1. From a periodic function to its sliding window point cloud. **Left:** A periodic function f. **Right:** Multidimensional scaling into $\mathbb{R}^3$ for $SW_{20,\tau}f$. For each t, we use the same color for $f(t)$ and $SW_{20,\tau}f(t)$. Please refer to an electronic version for colors.

Contributions

Apply persistent homology to study periodic data with phase variation, with guarantees.

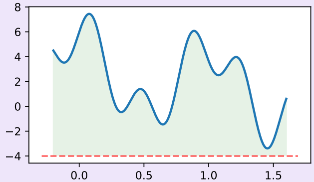
1. Describe the structure of a topological descriptor of S
2. Estimate $\gamma(T) - \gamma(0)$ and construct $\hat{\gamma}$ an estimator of γ.
3. Construct a descriptor of ϕ.

Outline

Persistence diagram of sub level sets

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t\in\mathbb{R}}$ as t increases.

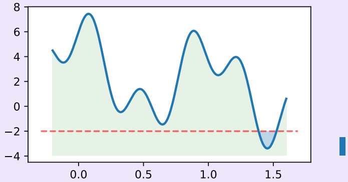


[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Persistence diagram of sub level sets

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t\in\mathbb{R}}$ as t increases.

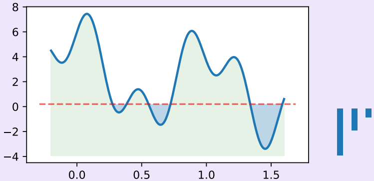


[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Persistence diagram of sub level sets

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t\in\mathbb{R}}$ as t increases.

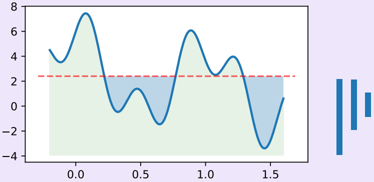


[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Persistence diagram of sub level sets

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t \in \mathbb{R}}$ as t increases.

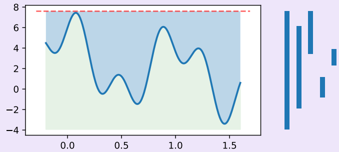


[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Persistence diagram of sub level sets

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t\in\mathbb{R}}$ as t increases.

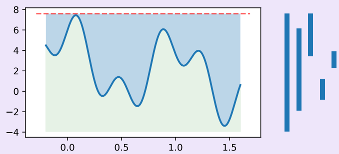


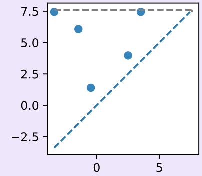
[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Persistence diagram of sub level sets

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t\in\mathbb{R}}$ as t increases.



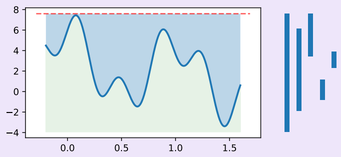


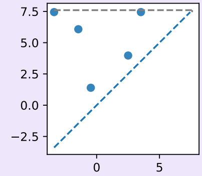
[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Persistence diagram of sub level sets

Intuition

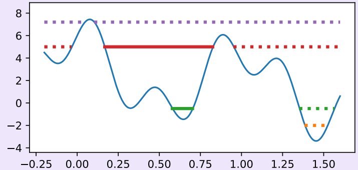
The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$, which reflect when connected components appear and merge in $(f^{-1}(]-\infty, t]))_{t \in \mathbb{R}}$ as t increases.

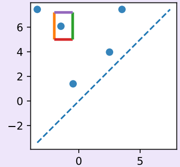




Definition

For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.





If $f : \mathbb{X} \to \mathbb{R}$ is continuous and $\mathbb{X}$ compact, then $D(f)$ is well-defined[1].

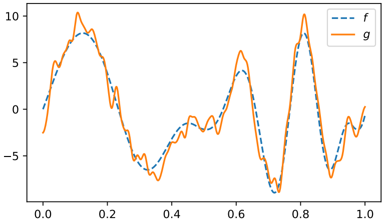
[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

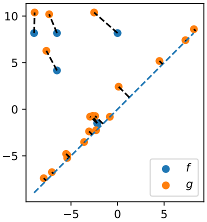
Stability: bottleneck distance

Theorem (Bottleneck stability of diagrams)

Let $f, g : \mathbb{X} \to \mathbb{R}$ be two continuous functions on a compact space $\mathbb{X}$. Then,

$$d_B(D(f), D(g)) \leq \|f - g\|_\infty.$$





Total p-persistence

Definition

The total p-persistence of a diagram D is

$$\mathrm{pers}_p(D) := \left(\sum_{(b,d)\in D} (d-b)^p \right)^{1/p} .$$

Proposition (Plonka and Zheng 2016, Perez 2022)

For $p = 1$, $\mathrm{pers}_1(D(f)) + \mathrm{pers}_1(D(-f)) = TV(f)$.
If f is α-Hölder for $p > 1 + 1/\alpha$, then, $\mathrm{pers}_p(D(f)) < \infty$.

[2] Olympio Hacquard et al. (2021). "Topologically Penalized Regression on Manifolds". In: *arXiv:2110.13749 [cs, math, stat]*. arXiv: `2110.13749 [cs, math, stat]`. (Visited on 01/04/2022), Section 3.2.

Total p-persistence

Definition

The total p-persistence of a diagram D is

$$\mathrm{pers}_p(D) := \left(\sum_{(b,d)\in D} (d-b)^p \right)^{1/p} .$$

Proposition (Plonka and Zheng 2016, Perez 2022)

For $p = 1$, $\mathrm{pers}_1(D(f)) + \mathrm{pers}_1(D(-f)) = TV(f)$. *(Only* dim $= 1$*!)*[2]
If f *is* α*-Hölder for* $p > 1 + 1/\alpha$*, then,* $\mathrm{pers}_p(D(f)) < \infty$*.*

[2] Olympio Hacquard et al. (2021). "Topologically Penalized Regression on Manifolds". In: *arXiv:2110.13749 [cs, math, stat]*. arXiv: 2110.13749 [cs, math, stat]. (Visited on 01/04/2022), Section 3.2.

Persistence diagrams of periodic functions

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function and denote by $\phi|_{[a,b]}$ the restriction of ϕ to an interval $[a, b]$.

Proposition (Invariance to reparametrisation)

Let $\gamma : [0, 1] \to [0, 1]$ be an increasing homeomorphism. Then, $D(\phi \circ \gamma) = D(\phi|_{[0,1]})$.

Theorem (Additivity of persistence diagrams for periodic functions)

For $R > 1$, there exists $c \in [0, 1]$ such that

$$D(\phi|_{[0,R]}) = \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D', \quad \textit{with} \ \ \mathrm{pers}_p(D') \leq 2\mathrm{pers}_p(D(\phi|_{[c,c+1]})). \tag{3}$$

Persistence diagrams of periodic functions

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function and denote by $\phi|_{[a,b]}$ the restriction of ϕ to an interval $[a, b]$.

Proposition (Invariance to reparametrisation)

Let $\gamma : [0, 1] \to [0, 1]$ be an increasing homeomorphism. Then, $D(\phi \circ \gamma) = D(\phi|_{[0,1]})$.

Theorem (Additivity of persistence diagrams for periodic functions)

For $R > 1$, there exists $c \in [0, 1]$ such that

$$D(\phi|_{[0,R]}) = \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D', \quad \textit{with} \ \ \mathrm{pers}_p(D') \leq 2\mathrm{pers}_p(D(\phi|_{[c,c+1]})). \tag{3}$$

If $R \in \mathbb{N}^*$, then

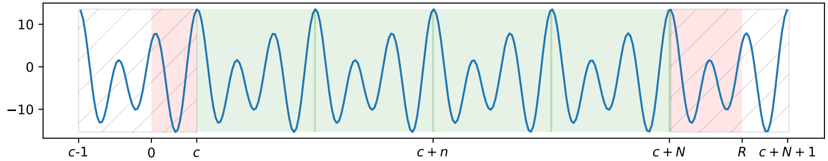
$$D(\phi|_{[0,R]}) = RD(\phi|_{[c,c+1]}). \tag{4}$$

Conclusion

The persistence diagram $D(\phi \circ \gamma)$ contains information about

- extrema of ϕ,
- number of periods $(\gamma(1) - \gamma(0))$.

Proof



Proof.

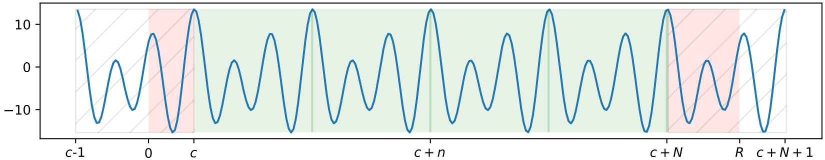
Let $c := \inf\{x \in [0,1[\mid \phi(x) = \max \phi\}$, $N = \max\{n \in \mathbb{N} \mid c + n \leq R\}$ and denote by $X_t := \phi^{-1}([-\infty, t[)$.

Step 1: For any $t < M$, $X_t \cap [0, c] \cap [c, c+1] = \emptyset$, so

$$H_0(X_t \cap [0, R]) \simeq H_0(X_t \cap [0, c]) \oplus \color{green}{H_0(X_t \cap [c, c+N])} \oplus H_0(X_t \cap [c+N, R])), \tag{5}$$

□

Proof



Proof.

Let $c := \inf\{x \in [0,1[\mid \phi(x) = \max\phi\}$, $N = \max\{n \in \mathbb{N} \mid c + n \leq R\}$ and denote by $X_t := \phi^{-1}([-\infty, t[)$.

Step 1: For any $t < M$, $X_t \cap [0,c] \cap [c, c+1] = \emptyset$, so

$$H_0(X_t \cap [0,R]) \simeq H_0(X_t \cap [0,c]) \oplus H_0(X_t \cap [c, c+N]) \oplus H_0(X_t \cap [c+N, R])), \tag{5}$$

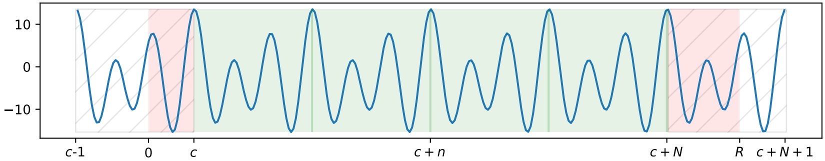
Step 2: similarly,

$$H_0(X_t \cap [c, c+N]) \simeq \bigsqcup_{n=1}^{N} H_0(X_t \cap [c + (n-1), c+n]) \tag{6}$$

$$(x \mapsto x + n) \quad \simeq \bigsqcup_{n=1}^{N} H_0(X_t \cap [c, c+1]) \tag{7}$$

$\square$

Proof



Proof.

Let $c := \inf\{x \in [0,1[|\ \phi(x) = \max \phi\}$, $N = \max\{n \in \mathbb{N} \mid c + n \leq R\}$ and denote by $X_t := \phi^{-1}([-\infty, t[)$.

Step 1: For any $t < M$, $X_t \cap [0, c] \cap [c, c+1] = \emptyset$, so

$$H_0(X_t \cap [0,R]) \simeq H_0(X_t \cap [0,c]) \oplus H_0(X_t \cap [c, c+N]) \oplus H_0(X_t \cap [c+N, R])), \tag{5}$$

Step 2: similarly,

$$H_0(X_t \cap [c, c+N]) \simeq \bigsqcup_{n=1}^{N} H_0(X_t \cap [c + (n-1), c+n]) \tag{6}$$

$$(x \mapsto x + n) \quad \simeq \bigsqcup_{n=1}^{N} H_0(X_t \cap [c, c+1]) \tag{7}$$

Step 3: The inclusion $[0, c] \subset [c-1, c]$ induces an injective morphism

$$H_0(X_t \cap [0, c]) \hookrightarrow H_0(X_t \cap [c-1, c]).$$

$\square$

Plan

Additivity of persistence diagrams of periodic functions

Segmentation of periodic signals

Signatures of periodic signals with phase variation

Segmenting a periodic curve with phase variation

Problem: phase estimation

Given S, estimate γ

$$\begin{array}{rccl} S: & [0,T] & \to & \mathbb{R} \\ & t & \mapsto & \phi(\gamma(t)) + W(t). \end{array}$$

Setting:

1. ϕ is unknown.
2. The number of periods N is an integer: $\gamma : [0, T] \to [0, N]$ with $N \in \mathbb{N}$.

Segmenting a periodic curve with phase variation

Problem: phase estimation

Given S, estimate γ

$$\begin{aligned} S: \quad [0,T] &\to \mathbb{R} \\ t &\mapsto \phi(\gamma(t)) + W(t). \end{aligned}$$

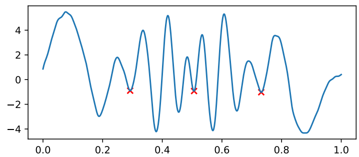
Setting:

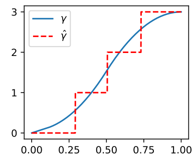
1. ϕ is unknown.
2. The number of periods N is an integer: $\gamma : [0,T] \to [0,N]$ with $N \in \mathbb{N}$.

Proposed solution: odometry

1. Estimate N.
2. Find $t_1 < \ldots < t_N$ such that $\gamma(t_n) - \gamma(t_{n-1}) = 1$ for all $n = 2, \ldots, N$.

Let $\hat{\gamma} : [0,T] \to \mathbb{R}^*$ be such that $\hat{\gamma}(t_n) = n$ and interpolate.





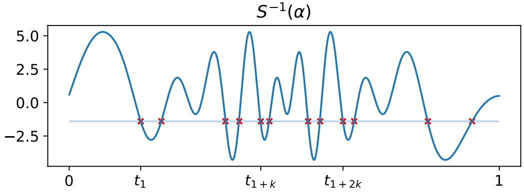
Developed and studied in (Bonis et al. 2022).

Zero-crossings

Let $K := |\phi^{-1}(\alpha) \cap [0, 1[|$ and assume that $0 < K < \infty$, for some $\alpha \in \mathbb{R}$.

Estimation of N

If K is known, $N_\alpha(S) := \frac{|S^{-1}(\alpha)|}{K}$ is an estimator of N.



Segmentation of the signal

If $S^{-1}(\alpha) = \{t_1, \ldots, t_{NK}\}$, then $\gamma(t_{n+k}) - \gamma(t_n) = 1$ for $1 \le k \le K$ and $n \le NK - k$.

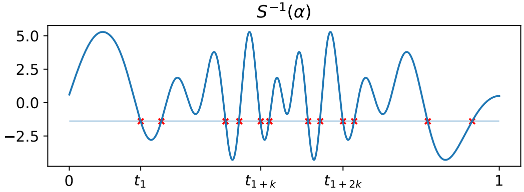
[3] Sunia Tanweer, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing.* arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

Zero-crossings

Let $K := |\phi^{-1}(\alpha) \cap [0, 1[|$ and assume that $0 < K < \infty$, for some $\alpha \in \mathbb{R}$.

Estimation of N

If K is known, $N_\alpha(S) := \frac{|S^{-1}(\alpha)|}{K}$ is an estimator of N.



Segmentation of the signal

If $S^{-1}(\alpha) = \{t_1, \dots, t_{NK}\}$, then $\gamma(t_{n+k}) - \gamma(t_n) = 1$ for $1 \le k \le K$ and $n \le NK - k$.

Issues

- K is not known (and not necessarily finite)
- N_α is not stable[3]

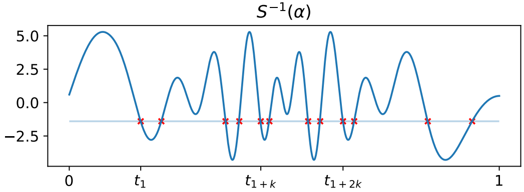
[3] Sunia Tanweer, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

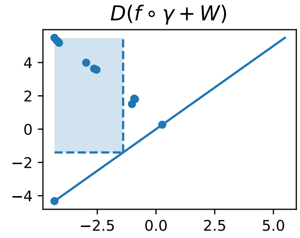
Zero-crossings

Let $K := |\phi^{-1}(\alpha) \cap [0,1[|$ and assume that $0 < K < \infty$, for some $\alpha \in \mathbb{R}$.

Estimation of N

If K is known, $N_\alpha(S) := \frac{|S^{-1}(\alpha)|}{K}$ is an estimator of N.





Segmentation of the signal

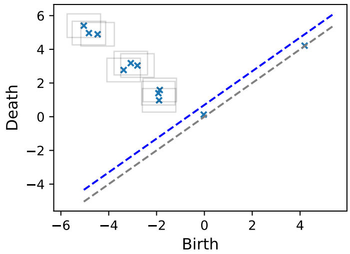
If $S^{-1}(\alpha) = \{t_1, \ldots, t_{NK}\}$, then $\gamma(t_{n+k}) - \gamma(t_n) = 1$ for $1 \leq k \leq K$ and $n \leq NK - k$.

Issues

- K is not known (and not necessarily finite)
- N_α is not stable[3]

[3] Sunia Tanweer, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

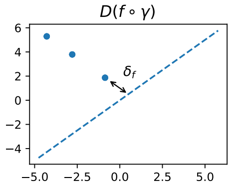
Estimation of N: stability



Estimator

For $\tau > 0$, we define

$$\hat{N}_\tau(S) := \gcd\{|D(S) \cap B(x, \tau)| \mid x \in D(S), \mathrm{pers}(x) > \tau\}. \tag{8}$$



Separation constant

$$\delta_\phi := \min(d(x_1, x_2), d(x_1, \Delta) \mid x_1, x_2 \in D(\phi))$$

Proposition (Stability)

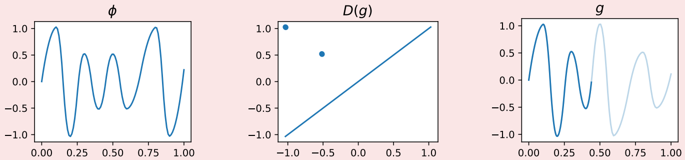
For $\tau > 0$ satisfying $2\|W\|_\infty < \tau < \delta/3$, we have that

$$\hat{N}_\tau(S) = \hat{N}_\tau(\phi \circ \gamma).$$

Correctness

Identifiability with the diagram

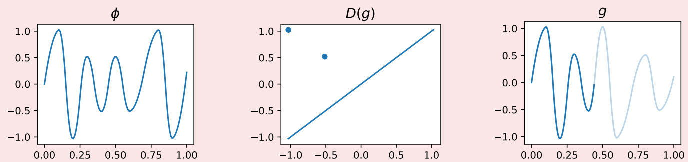
There exists a 1-periodic function g such that $D(g|_{[0,1]}) = D(\phi|_{[0,1]})/\hat{N}(\phi|_{[0,1]})$!



Correctness

Identifiability with the diagram

There exists a 1-periodic function g such that $D(g|_{[0,1]}) = D(\phi|_{[0,1]})/\hat{N}(\phi|_{[0,1]})$!



Non-degeneracy

We say that $\phi|_{[0,1]}$ is **non-degenerate** if $\hat{N}(\phi|_{[0,1]}) = 1$,

$$\hat{N}(\phi|_{[0,1]}) := \gcd \left\{ \lim_{\tau \to 0^+} |D(\phi|_{[0,1]}) \cap B(x,\tau)| \mid x \in D(\phi) \right\}. \tag{9}$$

Example

If ϕ has at least one unique critical value, it is non-degenerate.

Corollary (Stability)

If ϕ is non-degenerate, then for any $\tau > 0$ such that $2\|W\|_\infty < \tau < \delta/3$, we have

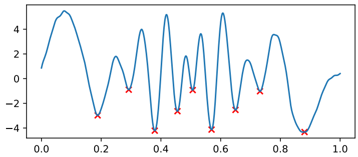
$$\hat{N}_\tau(S) = N.$$

Odometric sequence

Proposition

Let $\tau > 0$ and $\hat{\mathcal{C}}_\tau$ be the set of local minima of S, corresponding to points in the diagram with persistence more than τ. If $\tau \in]2\epsilon, \delta/3[$, then

$$|\hat{\mathcal{C}}_\tau| = NK.$$



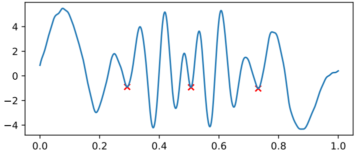
Put the figure with $K = 3$ sequences!

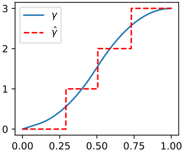
Odometric sequence

Proposition

Let $\tau > 0$ and $\hat{\mathcal{C}}_\tau$ be the set of local minima of S, corresponding to points in the diagram with persistence more than τ. If $\tau \in]2\epsilon, \delta/3[$, then

$$|\hat{\mathcal{C}}_\tau| = NK.$$





Put the figure with $K = 3$ sequences!

Conclusion and future work

Example

We can infer the number of periods of ϕ in S by counting points in the persistence diagram. We can also construct a segmentation of $[0, T]$ into periods.

Limitations and future work

1. Identifiability
2. More robust estimators
3. The method is applicable only to $N \in \mathbb{N}$: boundary effects when $\gamma(1) - \gamma(0) \notin \mathbb{N}$.

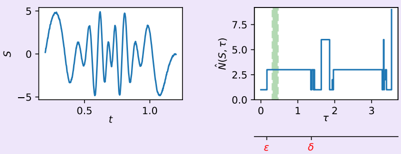
Conclusion and future work

Example

We can infer the number of periods of ϕ in S by counting points in the persistence diagram. We can also construct a segmentation of $[0, T]$ into periods.

Limitations and future work

1. Identifiability
   - Use merge trees to verify that the segmentation is correct
2. More robust estimators
   - Extend the guarantees to $\hat{N}_c$ and $\hat{N}_T$
   - Choose the sets to count multiplicity differently



3. The method is applicable only to $N \in \mathbb{N}$: boundary effects when $\gamma(1) - \gamma(0) \notin \mathbb{N}$.

Outline

Problem statement

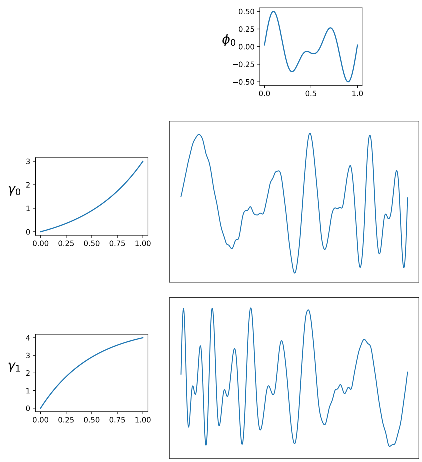
Data

Consider $S = \phi \circ \gamma + W$, where

- $\phi : \mathbb{R} \to \mathbb{R}$ is 1–periodic,
- $\gamma : [0, T] \to \mathbb{R}$ an increasing bijection,
- $W : [0, T] \to \mathbb{R}$ is a cont. stoch. proc.

Aim

Given S, construct a signature of ϕ.



Problem statement

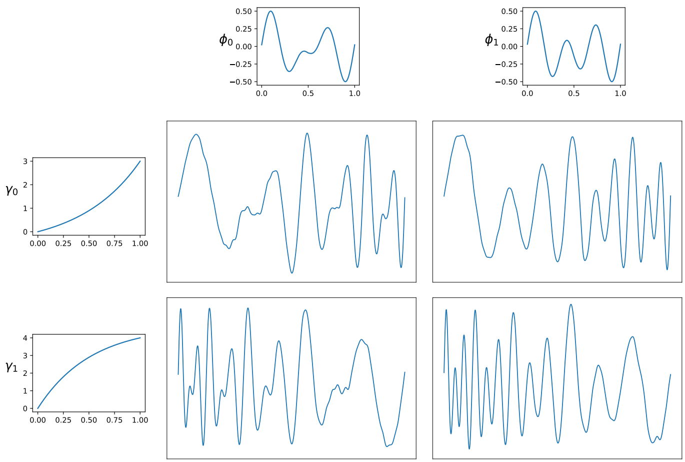
Data

Consider $S = \phi \circ \gamma + W$, where

- $\phi : \mathbb{R} \to \mathbb{R}$ is 1–periodic,
- $\gamma : [0, T] \to \mathbb{R}$ an increasing bijection,
- $W : [0, T] \to \mathbb{R}$ is a cont. stoch. proc.

Aim

Given S, construct a signature of ϕ.



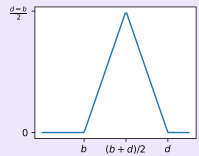
Functional representations of persistence diagrams

Persistence diagrams can be seen as multi-sets or discrete measures: even when a mean persistence diagram exists, it is not necessarily unique![4,5] It is common to represent a persistence diagram D in a functional space[6,7].

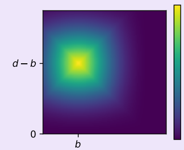
Functional representation

Let $\mathcal{H}$ be a functional Banach space

$$\begin{array}{rccccccc} \kappa: & \mathbb{R}^2 & \to & \mathcal{H} & & & & \\ & (b,d) & \mapsto & \kappa_{(b,d)}: & \mathbb{T} & \to & \mathbb{R} & \\ & & & & x & \mapsto & \kappa_{(b,d)}(x). & \end{array}$$



Persistence silhouette[8]



Persistence image[9]

Normalized functionals of persistence diagrams

For some $p \geq 1$ and $\epsilon > 0$,

$$\bar{\rho}(D) := \frac{\sum_{(b,d)\in D} w(d-b)\kappa_{(b,d)}}{\sum_{(b,d)\in D} w(d-b)}, \qquad \text{where } w(d-b) = \max(d-b-\epsilon, 0)^p. \tag{10}$$

[4] Yuriy Mileyko, Sayan Mukherjee, and John Harer (2011). "Probability Measures on the Space of Persistence Diagrams". In: *Inverse Problems* 27.12, p. 124007. ISSN: 0266-5611, 1361-6420. DOI: 10.1088/0266-5611/27/12/124007. (Visited on 08/18/2021).

[5] Vincent Divol and Théo Lacombe (2021). "Estimation and Quantization of Expected Persistence Diagrams". In: *arXiv:2105.04852 [math, stat]*. arXiv: 2105.04852 [math, stat]. (Visited on 05/31/2021).

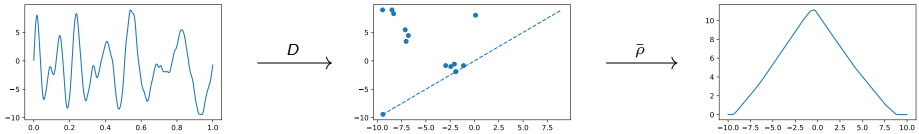
[6] Frédéric Chazal and Bertrand Michel (2021). "An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists". In: *Frontiers in Artificial Intelligence* 4. ISSN: 2624-8212. (Visited on 12/15/2022).

[7] Eric Berry et al. (2020). "Functional Summaries of Persistence Diagrams". In: *Journal of Applied and Computational Topology* 4.2, pp. 211–262. ISSN: 2367-1734. DOI: 10.1007/s41468-020-00048-w. (Visited on 08/24/2023).

[8] Peter Bubenik (2015). "Statistical Topological Data Analysis Using Persistence Landscapes". In: *Journal of Machine Learning Research* 16.1, pp. 77–102

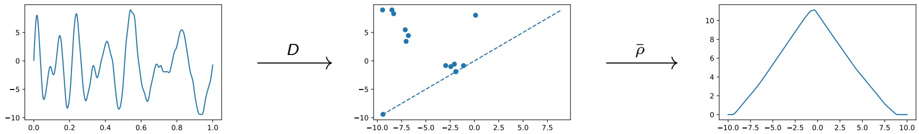
[9] Henry Adams et al. (2017). "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *The Journal of Machine Learning Research* 18.1, pp. 218–252

Proposed approach: normalized functionals of persistence



$$F(S) := \mathbb{E}[\bar{\rho}(D(S))]. \tag{11}$$

Proposed approach: normalized functionals of persistence



$$F(S) := \mathbb{E}[\bar{\rho}(D(S))]. \tag{11}$$

Properties

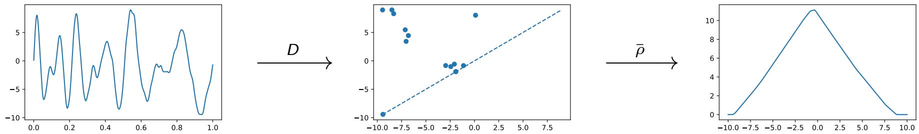
1. Consistency

$$\bar{\rho}(D(\phi|_{[0,R]})) \xrightarrow{\|\cdot\|_{\mathcal{H}}} \bar{\rho}(D(\phi|_{[c,c+1]})), \qquad \text{as } R \to \infty. \tag{12}$$

2. Stability

3. Estimation of a signature from a sampled time series

Proposed approach: normalized functionals of persistence



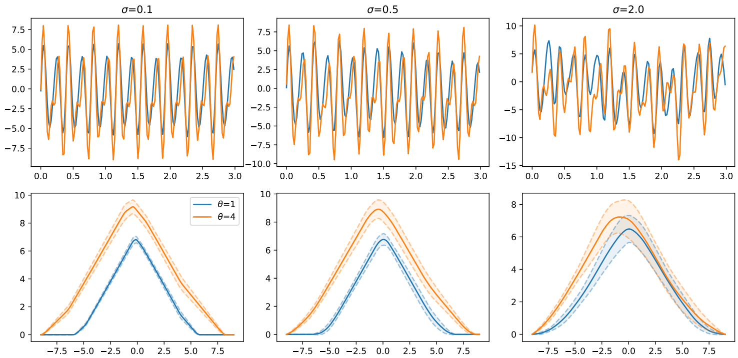
$$F(S) := \mathbb{E}[\bar{\rho}(D(S))]. \tag{11}$$

Properties

1. Consistency
$$\bar{\rho}(D(\phi|_{[0,R]})) \xrightarrow{\|\cdot\|_{\mathcal{H}}} \bar{\rho}(D(\phi|_{[c,c+1]})), \qquad \text{as } R \to \infty. \tag{12}$$
2. Stability
Let μ_1 and μ_2 be probability measures on a space of reparametrisations with fixed endpoints $\gamma(0), \gamma(T)$.
$$\|\mathbb{E}_{\gamma_1 \sim \mu_1, W}[\bar{\rho}(\phi \circ \gamma_1 + W)] - \mathbb{E}_{\gamma_2 \sim \mu_2, W}[\bar{\rho}(\phi \circ \gamma_2 + W)]\|_{\mathcal{H}} \leq C W_1(\mu_1, \mu_2)^{\alpha}. \tag{13}$$
3. Estimation of a signature from a sampled time series

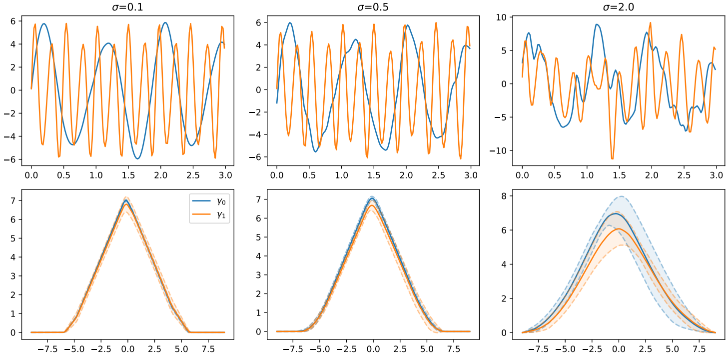
Numerical examples

$$\phi_1 \neq \phi_2, \qquad \mu_1 = \mu_2,$$



Numerical examples

$$\phi_1 = \phi_2, \qquad \mu_1 \neq \mu_2,$$



Estimation of signatures: introduction

Assume that only a single time series $\mathbf{S} = (S_n)_{n=1}^N \subset \mathbb{R}$ is given,

$$S_n = \phi(\gamma(t_n)) + W(t_n).$$

Can we estimate the signature of $\mathbf{S}$?

Proposition (Chazal et al. 2014[10], Berry et al. 2018[11])

Let be $D_1, \ldots, D_N$ i.i.d.persistence diagrams. When the (bracketing) entropy of $(\bar{\rho}_x)_{x \in \mathbb{T}}$ is finite,

$$\sqrt{N}\left(\frac{1}{N}\sum_{n=1}^{N}\bar{\rho}(D_n) - \bar{\rho}^*\right) \xrightarrow{d} \mathbb{G},$$

for a zero-mean stochastic process $\mathbb{G}$.

Challenges

- $\bar{\rho}$ is calculated on a window, not on a single element S_m
- $S_1, \ldots, S_N$ are not independent

[10] Frédéric Chazal et al. (2014). "Stochastic Convergence of Persistence Landscapes and Silhouettes". In: *Annual Symposium on Computational Geometry - SOCG'14*. Kyoto, Japan: ACM Press, pp. 474–483. ISBN: 978-1-4503-2594-3. DOI: 10.1145/2582112.2582128. (Visited on 03/05/2021)

[11] Eric Berry et al. (2018). *Functional Summaries of Persistence Diagrams*. arXiv: 1804.01618. (Visited on 03/21/2019)

Estimation of signatures: introduction

Assume that only a single time series $\mathbf{S} = (S_n)_{n=1}^{N} \subset \mathbb{R}$ is given,

$$S_n = \phi(\gamma(t_n)) + W(t_n).$$

Can we estimate the signature of $\mathbf{S}$?

Proposition (Chazal et al. 2014[10], Berry et al. 2018[11])

Let be $D_1, \ldots, D_N$ i.i.d.persistence diagrams. When the (bracketing) entropy of $(\bar{\rho}_x)_{x\in\mathbb{T}}$ is finite,

$$\sqrt{N}\left(\frac{1}{N}\sum_{n=1}^{N}\bar{\rho}(D_n) - \bar{\rho}^*\right) \xrightarrow{d} \mathbb{G},$$

for a zero-mean stochastic process $\mathbb{G}$.

Challenges

- $\bar{\rho}$ is calculated on a window, not on a single element S_m
  - $\bar{\rho}(\mathbf{S})$, where $\mathbf{S} := (S_1, \ldots, S_M)$ for some $M \in \mathbb{N}$
- $S_1, \ldots, S_N$ are not independent
  - Dependence in the window $\mathbf{S}$.
  - Dependence in γ_n and W_n.

[10] Frédéric Chazal et al. (2014). "Stochastic Convergence of Persistence Landscapes and Silhouettes". In: *Annual Symposium on Computational Geometry - SOCG'14*. Kyoto, Japan: ACM Press, pp. 474–483. ISBN: 978-1-4503-2594-3. DOI: 10.1145/2582112.2582128. (Visited on 03/05/2021)

[11] Eric Berry et al. (2018). *Functional Summaries of Persistence Diagrams*. arXiv: 1804.01618. (Visited on 03/21/2019)

Estimation of signatures: dependence

How does the dependence in both $(\gamma_n)_{n\in\mathbb{N}}$ and $(W_n)_{n\in\mathbb{N}}$ translate to that between $\mathbf{S}_1$ and $\mathbf{S}_n$?

Estimation of signatures: dependence

How does the dependence in both $(\gamma_n)_{n\in\mathbb{N}}$ and $(W_n)_{n\in\mathbb{N}}$ translate to that between $\mathbf{S}_1$ and $\mathbf{S}_n$?

Key insight

The following map is measurable

$$(\gamma_n, W_n) \mapsto (\phi(\gamma_n), W_n) \mapsto \phi(\gamma_n) + W_n, \qquad \text{and } \phi(\gamma_n) = \phi(\mathrm{frac}(\gamma_n)).$$

Estimation of signatures: dependence

How does the dependence in both $(\gamma_n)_{n\in\mathbb{N}}$ and $(W_n)_{n\in\mathbb{N}}$ translate to that between $\mathbf{S}_1$ and $\mathbf{S}_n$?

Key insight

The following map is measurable

$$(\gamma_n, W_n) \mapsto (\phi(\gamma_n), W_n) \mapsto \phi(\gamma_n) + W_n, \qquad \text{and } \phi(\gamma_n) = \phi(\mathrm{frac}(\gamma_n)).$$

Model for γ

For some $h > 0$, we set

$$\gamma_{n+1} = \gamma_n + hV_n,$$

for $(V_n)_{n\in\mathbb{N}}$ a stationary Markov chain on $[v_{\min}, v_{\max}] \subset]0, \infty[$. Then, $\mathrm{frac}(\gamma_n)$ is stationary and the strong-mixing coefficients of $(\gamma_n)_{n\in\mathbb{N}}$ decrease exponentially.

Estimation of signatures: dependence

How does the dependence in both $(\gamma_n)_{n\in\mathbb{N}}$ and $(W_n)_{n\in\mathbb{N}}$ translate to that between $\mathbf{S}_1$ and $\mathbf{S}_n$?

Key insight

The following map is measurable

$$(\gamma_n, W_n) \mapsto (\phi(\gamma_n), W_n) \mapsto \phi(\gamma_n) + W_n, \qquad \text{and } \phi(\gamma_n) = \phi(\mathrm{frac}(\gamma_n)).$$

Model for γ

For some $h > 0$, we set

$$\gamma_{n+1} = \gamma_n + hV_n,$$

for $(V_n)_{n\in\mathbb{N}}$ a stationary Markov chain on $[v_{\min}, v_{\max}] \subset]0, \infty[$. Then, $\mathrm{frac}(\gamma_n)$ is stationary and the strong-mixing coefficients of $(\gamma_n)_{n\in\mathbb{N}}$ decrease exponentially.

Idea of the proof

Show that for n big enough, some $c > 0$ and μ the uniform measure on $[0, 1]$,

$$P(\mathrm{frac}(\gamma_n) \in [a, b] \mid \gamma_0 = x_0) \geq c^n \mu([a, b]).$$

Estimation of signatures: procedure

Statistical model

A time series $(S_n)_{n=1}^N \subset \mathbb{R}$,

$$S_n = \phi(\gamma_n) + W_n, \tag{14}$$

- for $(V_n)_{n\in\mathbb{N}}$ a stationary Markov chain on $[v_{\min}, v_{\max}] \subset]0,\infty[$,

$$\gamma_{n+1} = \gamma_n + hV_n,$$

- $(W_n)_{n\in\mathbb{N}}$ a stationary, real-valued noise process, with strong-mixing coefficients $\beta_W(k) = O(k^{-3})$.

Procedure

1. Fix $M \in \mathbb{N}$,
2. Generate a sample $(\mathbf{S}_n)_{n=1}^{N-M-1}$, where $\mathbf{S}_n = (S_n, \dots, S_{n+M-1})$,
3. Calculate $\hat{F} := \frac{1}{N-M-1}\sum_{n=1}^{N-M-1} \bar{\rho}(\mathbf{S}_n)$.

Theorem

If the bracketing entropy of $(\bar{\rho}_x)_{x\in\mathbb{T}}$ is finite, then

$$\sqrt{N-M+1}(\hat{F} - \mathbb{E}[\bar{\rho}(\mathbf{S}_1)]) \to \mathbb{G} \tag{15}$$

where $\mathbb{G}$ is a zero–mean Gaussian process.

Numerical illustration

Hypothesis testing

Conclusion and future work

Conclusion

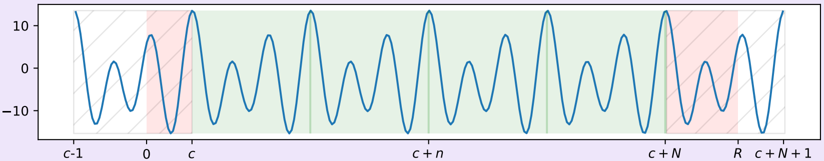
Normalized functionals of persistence of S yield **stable** signatures of ϕ and can be **estimated with standard techniques**.

Limitations and future work

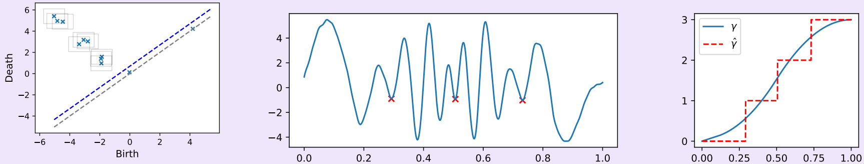
- Removing the assumption of fixed marginals from the stability
  - Understand the distance between $D(f|_{[0,T]})$ and $D(f|_{[0,t]}) \cup D(f|_{[t,T]})$
- Extensive numerical experiments and comparison with other methods.

Summary

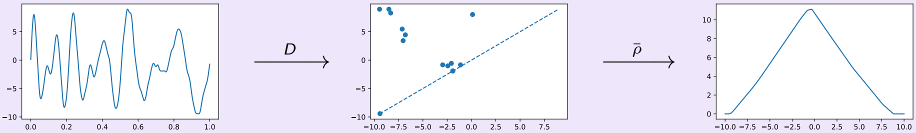
Additivity of diagrams



Odometry and phase estimation



Signatures



Refernces I

Adams, Henry et al. (2017). "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *The Journal of Machine Learning Research* 18.1, pp. 218–252.

Berry, Eric et al. (2018). *Functional Summaries of Persistence Diagrams*. arXiv: 1804.01618. (Visited on 03/21/2019).

– (2020). "Functional Summaries of Persistence Diagrams". In: *Journal of Applied and Computational Topology* 4.2, pp. 211–262. ISSN: 2367-1734. DOI: 10.1007/s41468-020-00048-w. (Visited on 08/24/2023).

Boashash, B. (1992). "Estimating and Interpreting the Instantaneous Frequency of a Signal. I. Fundamentals". In: *Proceedings of the IEEE* 80.4, pp. 520–538. ISSN: 1558-2256. DOI: 10.1109/5.135376.

Bois, Alexandre et al. (2022). "A Topological Data Analysis-Based Method for Gait Signals with an Application to the Study of Multiple Sclerosis". In: *PLOS ONE* 17.5. Ed. by Chan Hwang See, e0268475. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0268475. (Visited on 03/27/2023).

Bonis, Thomas et al. (2022). *Topological Phase Estimation Method for Reparameterized Periodic Functions*. DOI: 10.48550/arXiv.2205.14390. arXiv: 2205.14390 [cs, eess, math]. (Visited on 06/14/2022).

Bradley, Richard C. (2005). "Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions". In: *Probability Surveys* 2, pp. 107–144. DOI: 10.1214/154957805100000104. (Visited on 08/18/2022).

Bubenik, Peter (2015). "Statistical Topological Data Analysis Using Persistence Landscapes". In: *Journal of Machine Learning Research* 16.1, pp. 77–102.

Refernces II

Chazal, Frédéric and Bertrand Michel (2021). "An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists". In: *Frontiers in Artificial Intelligence* 4. ISSN: 2624-8212. (Visited on 12/15/2022).

Chazal, Frédéric et al. (2014). "Stochastic Convergence of Persistence Landscapes and Silhouettes". In: *Annual Symposium on Computational Geometry - SOCG'14*. Kyoto, Japan: ACM Press, pp. 474–483. ISBN: 978-1-4503-2594-3. DOI: 10.1145/2582112.2582128. (Visited on 03/05/2021).

Chazal, Frédéric et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Divol, Vincent and Théo Lacombe (2021). "Estimation and Quantization of Expected Persistence Diagrams". In: *arXiv:2105.04852 [math, stat]*. arXiv: 2105.04852 [math, stat]. (Visited on 05/31/2021).

Doukhan, Paul (1995). *Mixing*. Vol. 85. Lecture Notes in Statistics. Springer New York, NY. ISBN: 978-1-4612-2642-0. (Visited on 08/29/2022).

Hacquard, Olympio et al. (2021). "Topologically Penalized Regression on Manifolds". In: *arXiv:2110.13749 [cs, math, stat]*. arXiv: 2110.13749 [cs, math, stat]. (Visited on 01/04/2022).

Herbert Edelsbrunner and John Harer (2010). *Computational Topology: An Introduction*. American Mathematical Society. ISBN: 978-0-8218-4925-5.

Kennedy, Sean M., John D. Roth, and James W. Scrofani (2018). "A Novel Method for Topological Embedding of Time-Series Data". In: *2018 26th European Signal Processing Conference (EUSIPCO)*. Rome: IEEE, pp. 2350–2354. ISBN: 978-90-827970-1-5. DOI: 10.23919/EUSIPCO.2018.8553502. (Visited on 01/11/2021).

Refernces III

Khasawneh, Firas A. and Elizabeth Munch (2018). "Topological Data Analysis for True Step Detection in Periodic Piecewise Constant Signals". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 474.2218, p. 20180027. DOI: 10.1098/rspa.2018.0027. (Visited on 11/30/2020).

Mileyko, Yuriy, Sayan Mukherjee, and John Harer (2011). "Probability Measures on the Space of Persistence Diagrams". In: *Inverse Problems* 27.12, p. 124007. ISSN: 0266-5611, 1361-6420. DOI: 10.1088/0266-5611/27/12/124007. (Visited on 08/18/2021).

Perea, Jose A., Elizabeth Munch, and Firas A. Khasawneh (2019). "Approximating Continuous Functions on Persistence Diagrams Using Template Functions". In: *arXiv:1902.07190 [cs, math, stat]*. arXiv: 1902.07190 [cs, math, stat]. (Visited on 01/29/2020).

Perez, Daniel (2022). *On C0-persistent Homology and Trees*. DOI: 10.48550/arXiv.2012.02634. arXiv: 2012.02634v3.

Plonka, Gerlind and Yi Zheng (2016). "Relation between Total Variation and Persistence Distance and Its Application in Signal Processing". In: *Advances in Computational Mathematics* 42.3, pp. 651–674. ISSN: 1572-9044. DOI: 10.1007/s10444-015-9438-8. (Visited on 05/05/2021).

Srivastava, A et al. (2011). "Shape Analysis of Elastic Curves in Euclidean Spaces". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7, pp. 1415–1428. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2010.184. (Visited on 05/05/2021).

Su, Jingyong et al. (2014). "Statistical Analysis of Trajectories on Riemannian Manifolds: Bird Migration, Hurricane Tracking and Video Surveillance". In: *The Annals of Applied Statistics* 8.1, pp. 530–552. ISSN: 1932-6157. DOI: 10.1214/13-AOAS701. (Visited on 06/13/2020).

Refernces IV

Tang, R. and H.-G. Muller (2008). "Pairwise Curve Synchronization for Functional Data". In: *Biometrika* 95.4, pp. 875–889. ISSN: 0006-3444, 1464-3510. DOI: 10.1093/biomet/asn047. (Visited on 06/09/2023).

Tanweer, Sunia, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

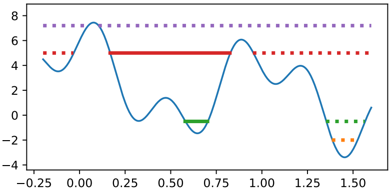
Zhao, Jiaping and Laurent Itti (2018). "shapeDTW: Shape Dynamic Time Warping". In: *Pattern Recognition* 74, pp. 171–184. ISSN: 00313203. DOI: 10.1016/j.patcog.2017.09.020. (Visited on 11/19/2020).

Thank you!

Persistence diagram of sub level sets: Definition

1. Persistence module

For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.



Persistence diagram of sub level sets: Definition

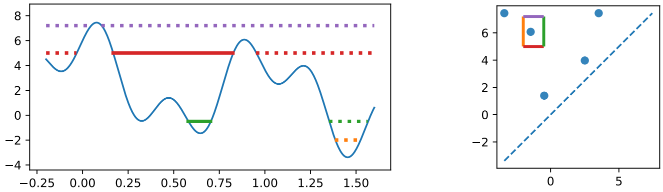
1. Persistence module

For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.

2. Rectangle measure

A measure m on rectangles of $\mathbb{R}^2$.

$$m(]a,b] \times [c,d[) = \dim\left(\frac{im(\iota_b^c) \cap \ker(\iota_c^d)}{im(\iota_a^c) \cap \ker(\iota_c^d)}\right),$$



Persistence diagram of sub level sets: Definition

1. Persistence module

For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.

2. Rectangle measure

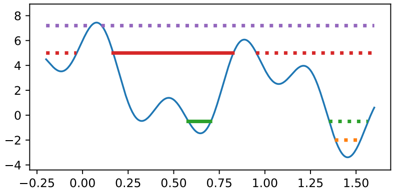
A measure m on rectangles of $\mathbb{R}^2$.

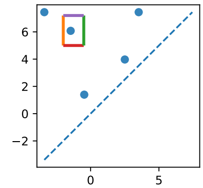
$$m(]a, b] \times [c, d[) = \dim \left(\frac{im(\iota_b^c) \cap \ker(\iota_c^d)}{im(\iota_a^c) \cap \ker(\iota_c^d)} \right),$$

3. Persistence diagram

The persistence diagram $D(f)$ is a multi-set in $\mathbb{R}^2$, where $(s, t) \in \mathbb{R}^2$ has multiplicity

$$m(s, t) = \lim_{\delta \to 0^+} m(]s - \delta, s + \delta] \times [t - \delta, t + \delta[).$$





Proof of additivity of sub level sets: details

Proof.

Let $c := \inf\{x \in [0,1[| \ \phi(x) = \max \phi\}$, $N = \max\{n \in \mathbb{N} \mid c + n \leq R\}$ and denote by $\mathbb{X}_t := \phi^{-1}([-\infty, t[)$.
Step 1: For any $t < M$, $\mathbb{X}_t \cap [0, c] \cap [c, c+1] = \emptyset$, so

$$H_0(\mathbb{X}_t \cap [0, R]) \simeq H_0(\mathbb{X}_t \cap [0, c]) \oplus H_0(\mathbb{X}_t \cap [c, c+N]) \oplus H_0(\mathbb{X}_t \cap [c+N, R])), \tag{16}$$

Step 2: similarly,

$$H_0(\mathbb{X}_t \cap [c, c+N]) \simeq \bigsqcup_{n=1}^{N} H_0(\mathbb{X}_t \cap [c+(n-1), c+n]) \tag{17}$$

$$(x \mapsto x+n) \quad \simeq \bigsqcup_{n=1}^{N} H_0(\mathbb{X}_t \cap [c, c+1]) \tag{18}$$

Step 3: The inclusion $[0, c] \subset [c-1, c]$ induces an injective morphism

$$H_0(\mathbb{X}_t \cap [0, c]) \hookrightarrow H_0(\mathbb{X}_t \cap [c-1, c]).$$

$\square$

Stability: bottleneck distance (detailed)

Definition (Herbert Edelsbrunner and John Harer 2010, p. VIII.2)

We call a ϵ-matching between two persistence diagrams D and D' a bijection $\Gamma : A \to A'$ between some subsets of $A \subset D$ and $A' \subset D'$, considered with multiplicity, if

$$\begin{aligned} d_\infty(a, \Gamma(a)) &\leq \epsilon, && \text{for any } a \in A, \\ d_\infty(a, \Delta) &\leq \epsilon, && \text{for any } a \in (D \setminus A) \cup (D' \setminus A'). \end{aligned}$$

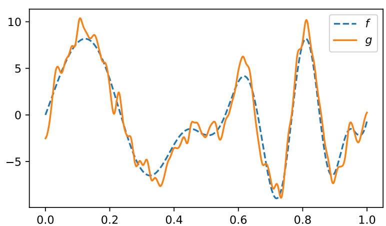
where $\Delta = \{(x, x) \in \mathbb{R}^2\}$ denotes the diagonal.

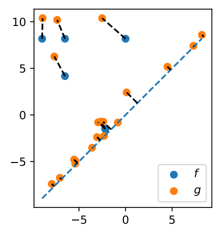
$$d_B(D, D') := \inf\{\epsilon > 0 \mid \Gamma \text{ is an } \epsilon\text{-matching between } D' \text{ and } D'\}.$$

Theorem (Bottleneck stability of diagrams)

Let $f, g : \mathbb{X} \to \mathbb{R}$ be two continuous functions on a compact space $\mathbb{X}$. Then,

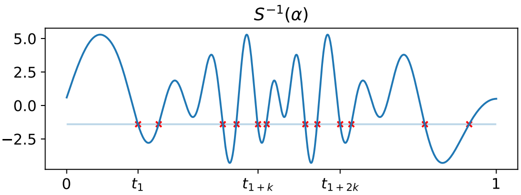
$$d_B(D(f), D(g)) \leq \|f - g\|_\infty.$$

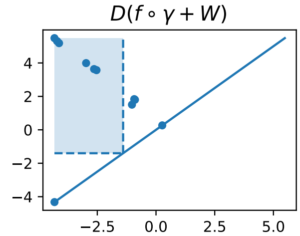




Zero-crossings from the persistence diagram

$$|S^{-1}(\alpha)| = 2 \lim_{\delta \to 0^+} |D(S) \cap (]-\infty, \alpha - \delta] \times [\alpha + \delta, \infty[)|.$$





Counting measure

The persistence diagram D is also a counting measure on rectangles $A \subset \Delta_+ = \{(b, d) \in \mathbb{R}^2 \mid x < y\}$. By (4),

$$|D(\phi \circ \gamma) \cap A| = N \, |D(\phi|_{[0,1]}) \cap A|$$

Application: magnetic odometry and speed estimation

Problem

Using the magnetic signal $\mathbf{B}$, recorded in a moving car, estimate the cars' trajectory. The angular position $t \mapsto \gamma(t)$ of a wheel in time is visible through $\mathbf{S}(t) = \mathbf{B}(\gamma(t), \gamma_h(t))$

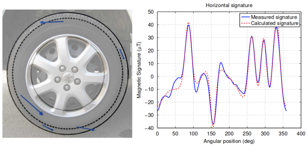


Fig. 2. Results of the optimization. Left: magnetic dipoles found after optimization. Right: measured and calculated signatures.

Proposed solution

1. $S := \langle \mathbf{S}, v \rangle$, project $\mathbf{S}$ along a suitable direction $v \in \mathbb{S}^2$
2. $\hat{N}_{c,\tau}(S)$, for an appropriate scale τ,
3. Derive an odometric sequence $t_1, \ldots, t_{\hat{N}_\tau(S)}$ from $\mathcal{C}_\tau$.
4. Construct $\hat{\gamma} : [0, T] \to \mathbb{R}$.

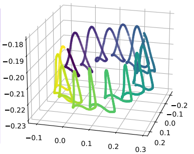
Results

Method	E_O $\mathbf{S}_{v_1}$	E_O $(\nabla \mathbf{S})_{v_1}$	E_I $\mathbf{S}_{v_1}$	E_I $(\nabla \mathbf{S})_{v_1}$
$\hat{N}_{c,\tau}$	15.75	16.66	3.02	3.01
$\hat{N}_{0,\tau}$	15.75	16.66	3.02	3.01
ZC	9.91	5.62	6.35	16.51

Application: magnetic odometry and speed estimation

Problem

Using the magnetic signal $\mathbf{B}$, recorded in a moving car, estimate the cars' trajectory. The angular position $t \mapsto \gamma(t)$ of a wheel in time is visible through $\mathbf{S}(t) = \mathbf{B}(\gamma(t), \gamma_h(t))$



Proposed solution

1. $S := \langle \mathbf{S}, v \rangle$, project $\mathbf{S}$ along a suitable direction $v \in \mathbb{S}^2$
2. $\hat{N}_{c,\tau}(S)$, for an appropriate scale τ,
3. Derive an odometric sequence $t_1, \ldots, t_{\hat{N}_\tau(S)}$ from $\mathcal{C}_\tau$.
4. Construct $\hat{\gamma} : [0, T] \to \mathbb{R}$.

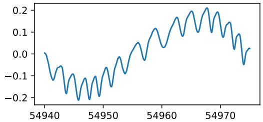
Results

	E_O		E_I	
Method	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$
$\hat{N}_{c,\tau}$	15.75	16.66	3.02	3.01
$\hat{N}_{0,\tau}$	15.75	16.66	3.02	3.01
ZC	9.91	5.62	6.35	16.51

Application: magnetic odometry and speed estimation

Problem

Using the magnetic signal $\mathbf{B}$, recorded in a moving car, estimate the cars' trajectory. The angular position $t \mapsto \gamma(t)$ of a wheel in time is visible through $\mathbf{S}(t) = \mathbf{B}(\gamma(t), \gamma_h(t))$



Proposed solution

1. $S := \langle \mathbf{S}, v \rangle$, project $\mathbf{S}$ along a suitable direction $v \in \mathbb{S}^2$
2. $\hat{N}_{c,\tau}(S)$, for an appropriate scale τ,
3. Derive an odometric sequence $t_1, \ldots, t_{\hat{N}_\tau(S)}$ from $\mathcal{C}_\tau$.
4. Construct $\hat{\gamma} : [0, T] \to \mathbb{R}$.

Results

	E_O		E_I	
Method	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$
$\hat{N}_{c,\tau}$	15.75	16.66	3.02	3.01
$\hat{N}_{0,\tau}$	15.75	16.66	3.02	3.01
ZC	9.91	5.62	6.35	16.51

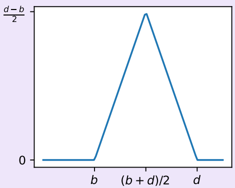
Normalized functionals of persistence

Functional representation

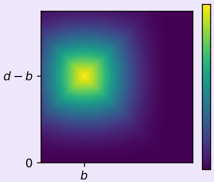
Let $\mathcal{H}$ be a functional Banach space

$$\begin{array}{rrcl} \kappa : & \mathbb{R}^2 & \to & \mathcal{H} \\ & (b,d) & \mapsto & \kappa_{(b,d)} : \begin{array}{rcl} \mathbb{T} & \to & \mathbb{R} \\ x & \mapsto & \kappa_{(b,d)}(x). \end{array} \end{array}$$

1. $\operatorname{supp}(\kappa_{(b,d)}) \subset K$, K bounded,
2. $x \mapsto \kappa_{(b,d)}(x)$ (uniformly) Lipschitz,
3. $\|\kappa_{(b,d)} - \kappa_{(b',d')}\|_{\mathcal{H}} \leq L_\kappa \|(b,d) - (b',d')\|$,
4. $\|\kappa_{(b,b)}\|_{\mathcal{H}} \leq C$.



Persistence silhouette[12]



Persistence image[13]

Normalized functionals of persistence diagrams

For some $p \geq 1$ and $\epsilon > 0$,

$$\bar{\rho}(D) := \frac{\sum_{(b,d)\in D} w(d-b)\kappa_{(b,d)}}{\sum_{(b,d)\in D} w(d-b)}, \qquad \text{where } w(d-b) = \max(d-b-\epsilon, 0)^p. \tag{19}$$

[12] Peter Bubenik (2015). "Statistical Topological Data Analysis Using Persistence Landscapes". In: *Journal of Machine Learning Research* 16.1, pp. 77–102

[13] Henry Adams et al. (2017). "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *The Journal of Machine Learning Research* 18.1, pp. 218–252

Measures of dependence

Types of dependence

There are different ways to measure dependence in a time series $(X_n)_{n\in\mathbb{N}} \subset \mathbb{X}$:

- m-dependence,
- strong-mixing,
- weak-dependence,

Strong mixing

The β-mixing coefficient of a time series $(X_n)_{n\in\mathbb{N}} \subset \mathbb{X}$ is

$$\beta_X(k) = \sup_{\mathcal{A},\mathcal{B}} \sum_{A\in\mathcal{A},B\in\mathcal{B}} |P(A\cap B) - P(A)P(B)|,$$

where $\mathcal{A} \subset \sigma^X_{-\infty,0}$, $\mathcal{B} \subset \sigma^X_{k,\infty}$ are finite partitions of the sample space and $\sigma^X_{a,b} := \sigma((X_n)_{a\leq n\leq b})$.

Example

1. If $(X_n)_n$ is m-dependent, then $\beta_X(k) = 0$ for $k \geq m$.
2. Markov chains: irreducible and aperiodic.

Proposition

For any measurable function $f : \mathbb{X} \to \mathbb{Y}$, $\beta_X(k) \leq \beta_{f(X)}(k)$.

Measures of dependence

Types of dependence

There are different ways to measure dependence in a time series $(X_n)_{n\in\mathbb{N}} \subset \mathbb{X}$:

- m-dependence,
- strong-mixing, $\longrightarrow$ **preserved by measurable functions!**
- weak-dependence,

Strong mixing

The β-mixing coefficient of a time series $(X_n)_{n\in\mathbb{N}} \subset \mathbb{X}$ is

$$\beta_X(k) = \sup_{\mathcal{A},\mathcal{B}} \sum_{A\in\mathcal{A},B\in\mathcal{B}} |P(A\cap B) - P(A)P(B)|,$$

where $\mathcal{A} \subset \sigma^X_{-\infty,0}$, $\mathcal{B} \subset \sigma^X_{k,\infty}$ are finite partitions of the sample space and $\sigma^X_{a,b} := \sigma((X_n)_{a\leq n\leq b})$.

Example

1. If $(X_n)_n$ is m-dependent, then $\beta_X(k) = 0$ for $k \geq m$.
2. Markov chains: irreducible and aperiodic.

Proposition

For any measurable function $f : \mathbb{X} \to \mathbb{Y}$, $\beta_X(k) \leq \beta_{f(X)}(k)$.