Time warping invariants and quasisymmetric functions

Joscha Diehl (Universität Greifswald)

joint with K. Ebrahimi-Fard (NTNU), N. Tapia (TU Berlin)

28. April - NASO

Slides at: `https://diehlj.github.io`

I hope to answer:

- What are time warping invariants?
- What (and why) are quasisymmetric functions?

To pique interest in the two people I know are in the audience (the other speakers)

- There is a map from matroids into quasisymmetric functions.[1] Is there a matroid structure visible in this talk? Does it matter for what we do?
- What is the *strength* of quasisymmetric polynomials/functions? Does it matter for what we do?

[1] *Billera, L.J., Jia, N. and Reiner, V., 2009. "A quasisymmetric function for matroids." European Journal of Combinatorics, 30(8), pp.1727-1757.*

I hope to answer:

- What are time warping invariants?
- What (and why) are quasisymmetric functions?

To pique interest in the two people I know are in the audience (the other speakers)

- There is a map from matroids into quasisymmetric functions.[1] Is there a matroid structure visible in this talk? Does it matter for what we do?
- What is the *strength* of quasisymmetric polynomials/functions? Does it matter for what we do?

[1] *Billera, L.J., Jia, N. and Reiner, V., 2009. "A quasisymmetric function for matroids." European Journal of Combinatorics, 30(8), pp.1727-1757.*

Time warping invariants

Often the speed at which a time series is run is unknown or does not matter.

Example

GPS measurements of people running around a lake. (1D-Example)

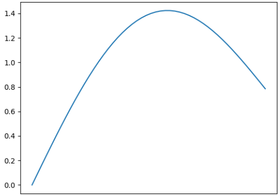


Fig: Runner A

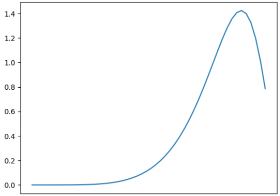


Fig: Runner B

They are both running the same route, just at different speeds.

Examples

- Heartbeat
- Different sampling speeds.

Continuous function

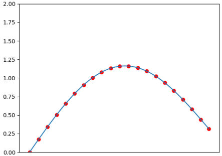


Fig: Slow sampling.

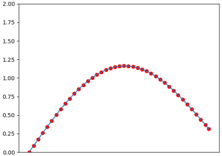
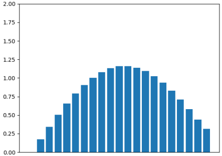
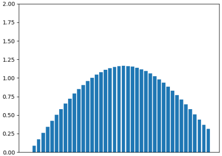


Fig: Fast sampling.

Discretely sampled





Invariance to speed is usually dealt with via the
dynamic time warping (DTW) distance
(immense amount of literature: Sakoe, Chiba, Keogh, Berndt, Clifford, Petitjean, Salvador, Chan, Müller, Jain, . . .)

This talk:
features that are invariant to time warping. (.. to be defined ..)

Mathematical formulation

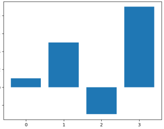
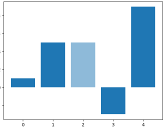
For us a time series is just a sequence of real numbers

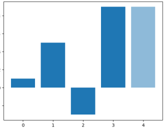
$$x = (x_1, x_2, .., x_{N-1}, x_N) \in \mathbb{R}^N.$$

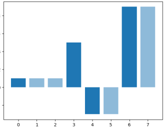
For convenience we work with an infinite time horizon and sequences that are eventually constant, i.e.

$$x = (x_1, x_2, .., x_{N-1}, x_N, x_N, x_N, \dots) \in \mathbb{R}^{\mathbb{N}}_{\text{ev. const.}}.$$

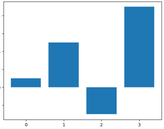
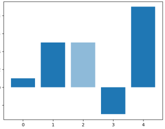
We say a function $F : \mathbb{R}^{\mathbb{N}}_{\text{ev. const.}} \to \mathbb{R}$ is time warping invariant (or *invariant to stuttering*) if, for example,

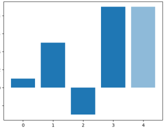
$F\Big($  $\Big) = F\Big($  $\Big)$

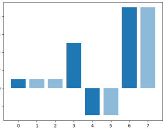
$= F\Big($  $\Big)$

$= F\Big($  $\Big)$.

We say a function $F : \mathbb{R}^{\mathbb{N}}_{\text{ev. const.}} \to \mathbb{R}$ is **time warping invariant** (or *invariant to stuttering*) if, for example,

$F\Big($  $\Big) = F\Big($  $\Big)$

$= F\Big($  $\Big)$

$= F\Big($  $\Big).$

What functions F can we cook up, that are invariant in this sense?

What functions F can we cook up, that are invariant in this sense?

$$F(x) := 17$$

Not very descriptive ...

$$F(x) := \sup_{n \leq N} x_n$$

$$F(x) := x_N - x_0 = \sum_{n \leq N} x_n - x_{n-1} =: \sum_{n \leq N} \delta x_n.$$

What functions F can we cook up, that are invariant in this sense?

$$F(x) := 17$$ *Not very descriptive . . .*

$$F(x) := \sup_{n \le N} x_n$$

$$F(x) := x_N - x_0 = \sum_{n \le N} x_n - x_{n-1} =: \sum_{n \le N} \delta x_n.$$

What functions F can we cook up, that are invariant in this sense?

$$F(x) := 17$$

Not very descriptive ...

$$F(x) := \sup_{n \leq N} x_n$$

$$F(x) := x_N - x_0 = \sum_{n \leq N} x_n - x_{n-1} =: \sum_{n \leq N} \delta x_n.$$

$$F(x) := x_N - x_0 = \sum_{n \leq N} x_n - x_{n-1} =: \sum_{n \leq N} \delta x_n.$$

This works for any $p = 1, 2, 3, ..$:

$$F(x) = \sum_{n \leq N} (\delta x_n)^p.$$

These are the (power-sum) symmetric polynomials in the increments.

Since we work with time-series that are eventually constant, we can write

$$F(x) = \sum_{n} (\delta x_n)^p,$$

the (power-sum) symmetric functions in the increments.

$$F(x) := x_N - x_0 = \sum_{n \le N} x_n - x_{n-1} =: \sum_{n \le N} \delta x_n.$$

This works for any $p = 1, 2, 3, ..$:

$$F(x) = \sum_{n \le N} (\delta x_n)^p.$$

These are the (power-sum) symmetric polynomials in the increments.

Since we work with time-series that are eventually constant, we can write

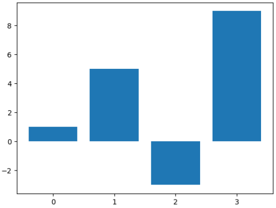
$$F(x) = \sum_{n} (\delta x_n)^p,$$

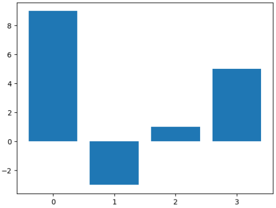
the (power-sum) symmetric functions in the increments.

$$F(x) = \sum_n (\delta x_n)^p.$$

The (power-sum) **symmetric functions** in the increments.

Clearly, these can <u>not</u> distinguish these two time series:





How to get the order of events ? Let us try

$$F(x) := \sum_{i_1 < i_2} (\delta x_{i_1})^2 \delta x_{i_2}.$$

This is not a symmetric function. Nonetheless:

$$F\Big((x_0, x_1, x_2, x_3, x_4, \dots)\Big)$$

$$= (\delta x_1)^2 \cdot \delta x_2$$

$$+ \Big[(\delta x_1)^2 + (\delta x_2)^2\Big] \cdot \delta x_3$$

$$+ \Big[(\delta x_1)^2 + (\delta x_2)^2 + (\delta x_3)^2\Big] \cdot \delta x_4$$

$$+ \dots$$

$$F\Big(z := (x_0, x_1, x_1, x_2, x_3, \dots)\Big)$$

$$= \underbrace{(\delta z_1)^2}_{(\delta x_1)^2} \cdot \underbrace{\delta z_2}_{=0}$$

$$+ \underbrace{\Big[(\delta z_1)^2 + (\delta z_2)^2\Big]}_{(\delta x_1)^2} \cdot \underbrace{\delta z_3}_{\delta x_2}$$

$$+ \underbrace{\Big[(\delta z_1)^2 + (\delta z_2)^2 + (\delta z_3)^2\Big]}_{(\delta x_1)^2 + (\delta x_2)^2} \cdot \underbrace{\delta z_4}_{\delta x_3}$$

$$+ ..$$

$\Rightarrow F$ is time warping invariant.

Along the same vain, all expressions of the form

$$\sum_{i_1 < \cdots < i_n} (\delta x_{i_1})^{\alpha_1} \ldots (\delta x_{i_n})^{\alpha_n},$$

are invariant to time warping.

QUESTIONS

- **A** Are these all of the (polynomial) invariants?
- **B** There are algebraic dependencies, like

$$\left(\sum_i \delta x_i\right)^2 = 2 \sum_{i_1 < i_2} \delta x_{i_1} \delta x_{i_2} + \sum_i (\delta x_i)^2.$$

How to disentangle them?

Comments at this point?

Along the same vain, all expressions of the form

$$\sum_{i_1 < \cdots < i_n} (\delta x_{i_1})^{\alpha_1} \ldots (\delta x_{i_n})^{\alpha_n},$$

are invariant to time warping.

QUESTIONS

- **A** Are these all of the (polynomial) invariants?
- **B** There are algebraic dependencies, like

$$\left(\sum_i \delta x_i\right)^2 = 2 \sum_{i_1 < i_2} \delta x_{i_1} \delta x_{i_2} + \sum_i (\delta x_i)^2.$$

How to disentangle them?

Comments at this point?

Quasisymmetric functions A remark on nomenclature.

These are far from generalizations for generalization's sake: they are most important objects in their own right [1].

[1]But it must be admitted that the name 'quasisymmetric functions' is suspicious. Practically all things 'quasi' have the taint of some largely useless generalization about them.

Hazewinkel, Gubareni, Kirichenko. "Algebras, Rings and Modules - Lie Algebras and Hopf Algebras", AMS, 2010.

Def A formal power series in commuting $X_1, X_2, X_3, ..$ is a quasisymmetric function if for all $i_1 < \cdots < i_n$, $j_1 < \cdots < j_n$ and all $\alpha_1, \ldots, \alpha_n \geq 1$ the cofficients of these monomials coincide,

$$X_{i_1}^{\alpha_1} \ldots X_{i_n}^{\alpha_n} \quad \text{and} \quad X_{j_1}^{\alpha_1} \ldots X_{j_n}^{\alpha_n}.$$

Examples

$$\sum_i X_i \qquad \sum_i (X_i)^2 \qquad \sum_{i_1 < i_2} X_{i_1} X_{i_2} \qquad \sum_{i_1 < i_2} (X_{i_1})^2 X_{i_2} \qquad \ldots$$

Quasisymmetric functions A remark on nomenclature.

> These are far from generalizations for generalization's sake: they are most important objects in their own right [1].
>
> [1]But it must be admitted that the name 'quasisymmetric functions' is suspicious. Practically all things 'quasi' have the taint of some largely useless generalization about them.

Hazewinkel, Gubareni, Kirichenko. "Algebras, Rings and Modules - Lie Algebras and Hopf Algebras", AMS, 2010.

Def A formal power series in commuting $X_1, X_2, X_3, ..$ is a quasisymmetric function if for all $i_1 < \cdots < i_n$, $j_1 < \cdots < j_n$ and all $\alpha_1, \ldots, \alpha_n \geq 1$ the cofficients of these monomials coincide,

$$X_{i_1}^{\alpha_1} \ldots X_{i_n}^{\alpha_n} \qquad \text{and} \qquad X_{j_1}^{\alpha_1} \ldots X_{j_n}^{\alpha_n}.$$

Examples

$$\sum_i X_i \qquad \sum_i (X_i)^2 \qquad \sum_{i_1 < i_2} X_{i_1} X_{i_2} \qquad \sum_{i_1 < i_2} (X_{i_1})^2 X_{i_2} \qquad \ldots$$

Lemma (D,Ebrahimi-Fard,Tapia '19)

All polynomial invariants to time warping are given by quasisymmetric functions (evaluated at the increments)

$$\sum_{i_1<\cdots<i_n} (\delta x_{i_1})^{\alpha_1} \cdot \cdots \cdot (\delta x_{i_n})^{\alpha_n}, \quad n \geq 1, \alpha \in \mathbb{N}^n_{\geq 1}.$$

This answers question A.

To answer question B, it is convenient to store all these invariants as a character on a (Hopf) algebra.[2]

[1]The Hopf algebra of quasisymmetric function was studied by *Malvenuto/Reutenauer 1994*.

Lemma (D,Ebrahimi-Fard,Tapia '19)

All polynomial invariants to time warping are given by quasisymmetric functions (evaluated at the increments)

$$\sum_{i_1 < \cdots < i_n} (\delta x_{i_1})^{\alpha_1} \cdot \cdots \cdot (\delta x_{i_n})^{\alpha_n}, \quad n \geq 1, \alpha \in \mathbb{N}^n_{\geq 1}.$$

This answers question A.

To answer question B, it is convenient to store all these invariants as a character on a (Hopf) algebra. [2]

---

[1]The Hopf algebra of quasisymmetric function was studied by *Malvenuto/Reutenauer 1994*.

Define the iterated-sums signature $(y := \delta x)$

$$
\begin{aligned}
\mathrm{ISS}(x) := 1 &+ \sum_{i_1} y_{i_1} \cdot 1 + \sum_{i_1} y_{i_1}^2 \cdot 2 + \sum_{i_1} y_{i_1}^3 \cdot 3 + \ldots \\
&+ \sum_{i_1 < i_2} y_{i_1} y_{i_2} \cdot 1|1 + \sum_{i_1 < i_2} y_{i_1}^2 y_{i_2} \cdot 2|1 + \ldots \\
&\in H := \mathbb{R} \oplus \mathbb{R}^{\mathbb{N}} \oplus (\mathbb{R}^{\mathbb{N}})^{\otimes 2} \oplus (\mathbb{R}^{\mathbb{N}})^{\otimes 3} \oplus {}^{3}..
\end{aligned}
$$

$$
\Big\langle 573, \mathrm{ISS}(x) \Big\rangle = \sum_{i_1 < i_2 < i_3} y_{i_1}^5 y_{i_2}^5 y_{i_3}^3 .
$$

In words: H is the space of formal (infinite) $\mathbb{R}$-linear combinations of words in $1, 2, \ldots$.
The product on H is concatenation of words, e.g.

$$
5|7 \bullet 3|1|2 = 5|7|3|1|2.
$$

[1] *Actually: direct product, not direct sum.*

Define the iterated-sums signature $(y := \delta x)$

$$\begin{aligned}\mathrm{ISS}(x) &:= 1 + \sum_{i_1} y_{i_1} \cdot 1 + \sum_{i_1} y_{i_1}^2 \cdot 2 + \sum_{i_1} y_{i_1}^3 \cdot 3 + \dots \\ &\quad + \sum_{i_1<i_2} y_{i_1} y_{i_2} \cdot 1|1 + \sum_{i_1<i_2} y_{i_1}^2 y_{i_2} \cdot 2|1 + \dots \\ &\in H := \mathbb{R} \oplus \mathbb{R}^{\mathbb{N}} \oplus (\mathbb{R}^{\mathbb{N}})^{\otimes 2} \oplus (\mathbb{R}^{\mathbb{N}})^{\otimes 3} \oplus {}^{3} ..\end{aligned}$$

$$\Big\langle 573, \mathrm{ISS}(x) \Big\rangle = \sum_{i_1<i_2<i_3} y_{i_1}^5 y_{i_2}^5 y_{i_3}^3.$$

In words: H is the space of formal (infinite) $\mathbb{R}$-linear combinations of words in $1, 2, \dots$.
The product on H is concatenation of words, e.g.

$$5|7 \bullet 3|1|2 = 5|7|3|1|2.$$

NONCOMMUTATIVE

[1] *Actually: direct product, not direct sum.*

Lemma (D,Ebrahimi-Fard,Tapia '19)

For $x, x' \in \mathbb{R}^{\mathbb{N}}_{ev.\ const.}$ let denote $x \sqcup x' \in \mathbb{R}^{\mathbb{N}}_{ev.\ const.}$ their concatenation. Then:

$$\mathrm{ISS}(x \sqcup x') = \mathrm{ISS}(x) \bullet \mathrm{ISS}(x') \qquad \text{"Chen's identity".}$$

Remark

This leads to a (naive) implementation, since ISS('one increment') is easy to calculate.

Lemma (D,Ebrahimi-Fard,Tapia '19)

For $x, x' \in \mathbb{R}^{\mathbb{N}}_{ev.\ const.}$ let denote $x \sqcup x' \in \mathbb{R}^{\mathbb{N}}_{ev.\ const.}$ their concatenation. Then:

$$\mathrm{ISS}(x \sqcup x') = \mathrm{ISS}(x) \bullet \mathrm{ISS}(x') \qquad \text{“\textit{Chen's identity}”.}$$

Remark

This leads to a (naive) implementation, since ISS('one increment') is easy to calculate.

Lemma (Malvenuto, Reutenauer '94; DEFT '19)

$$\Big\langle \alpha, \mathrm{ISS}(x) \Big\rangle \cdot \Big\langle \beta, \mathrm{ISS}(x) \Big\rangle = \Big\langle \alpha \sqcup\!\sqcup\, \beta, \mathrm{ISS}(x) \Big\rangle \quad \text{“shuffle identity”.}$$

Here $\sqcup\!\sqcup$ is the quasishuffle on H^*. For example

$$1 \sqcup\!\sqcup\, 1 = 2\ 1|1 + 2 \qquad \leftrightarrow \left(\sum_i \delta x_i \right)^2 = 2 \sum_{i_1 < i_2} \delta x_{i_1} \delta x_{i_2} + \sum_i (\delta x_i)^2.$$

$$1|2 \sqcup\!\sqcup\, 3 = 1|2|3 + 1|3|2 + 3|1|2 + 1|5 + 4|2 = 3 \sqcup\!\sqcup\, 1|2.$$

Lemma (Malvenuto, Reutenauer '94; DEFT '19)

$$\Big\langle \alpha, \mathsf{ISS}(x) \Big\rangle \cdot \Big\langle \beta, \mathsf{ISS}(x) \Big\rangle = \Big\langle \alpha \sqcup\!\sqcup\, \beta, \mathsf{ISS}(x) \Big\rangle$$ *"shuffle identity".*

Here $\sqcup\!\sqcup$ is the quasishuffle on H^*. For example

$$1 \sqcup\!\sqcup\, 1 = 2\ 1|1 + 2 \qquad \leftrightarrow \left(\sum_i \delta x_i \right)^2 = 2 \sum_{i_1 < i_2} \delta x_{i_1} \delta x_{i_2} + \sum_i (\delta x_i)^2.$$

COMMUTATIVE

$$1|2 \sqcup\!\sqcup\, 3 = 1|2|3 + 1|3|2 + 3|1|2 + 1|5 + 4|2 = 3 \sqcup\!\sqcup\, 1|2.$$

Lemma (Malvenuto, Reutenauer '94; DEFT '19)

$$\Big\langle \alpha, \mathrm{ISS}(x) \Big\rangle \cdot \Big\langle \beta, \mathrm{ISS}(x) \Big\rangle = \Big\langle \alpha \sqcup\!\sqcup \beta, \mathrm{ISS}(x) \Big\rangle$$ *"shuffle identity".*

Here $\sqcup\!\sqcup$ is the quasishuffle on H^*. For example

$$1 \sqcup\!\sqcup 1 = 2\ 1|1 + 2 \qquad \leftrightarrow \left(\sum_i \delta x_i \right)^2 = 2 \sum_{i_1 < i_2} \delta x_{i_1} \delta x_{i_2} + \sum_i (\delta x_i)^2.$$

COMMUTATIVE

$$1|2 \sqcup\!\sqcup 3 = 1|2|3 + 1|3|2 + 3|1|2 + 1|5 + 4|2 = 3 \sqcup\!\sqcup 1|2.$$

Just as for the iterated-integrals signature (Chen, Lyons, Reutenauer, ..), we get

Theorem (D,Ebrahimi-Fard,Tapia '19)

The iterated-sums signature is a character on the Hopf algebra, endowed with quasishuffle product and deconcatenation coproduct.

And this works for multidimensional time-series analogously.

Without going into details, this result gives that $\mathrm{ISS}(x)$ is grouplike, and in bijection to its logarithm

$$\log \mathrm{ISS}(x) = \mathrm{ISS}(x) + \frac{1}{2!}\mathrm{ISS}(x)^{\bullet 2} + \frac{1}{3!}\mathrm{ISS}(x)^{\bullet 3} + \dots$$

This logarithm lives in a Lie algebra and should yield the disentanglement of Question B.

	$\dim \mathrm{ISS}$ of weight n	$\dim \log \mathrm{ISS}$ of weight n
$n = 1$	1	1
$n = 2$	2	1
$n = 3$	3	2
$n = 4$	4	2

To really show that this is the *minimal* amount of information contained in $\mathrm{ISS}(x)$, one step is still missing though ... (Chow's theorem)

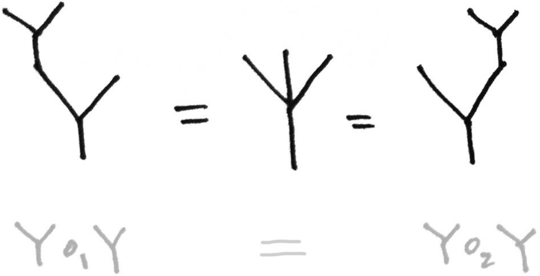
SUMMARY

- Time warping invariants are *very* relevant for time series analysis.
- All *polynomial* invariants are given by (evaluation of) quasisymmetric functions.
- They are conveniently stored in a Hopf algebra framework and provide a concrete view on these abstract structures.

THANK YOU!

ADVERTISEMENT

Noémie Combe (MPI Leipzig) and I are organizing an online reading group on operads .



If you are interested in joining: please get in touch!