An algebraic geometry of paths via the iterated-integrals signature

Rosa Lili Dora Preiß

MPI MiS Leipzig

Nonlinear algebra seminar, March 20 2024

Let $X : [0,T] \to \mathbb{R}^d$ be a continuous bounded variation path. We define the k-th *signature tensor* by

$$\sigma^{(k)}(X) := \int_{0<t_1<\cdots<t_k<T} X'(t_1) \otimes \cdots \otimes X'(t_k) \, \mathrm{d}t_1 \cdots \mathrm{d}t_k \in (\mathbb{R}^d)^{\otimes k}$$

Let $X : [0,T] \to \mathbb{R}^d$ be a continuous bounded variation path. We define the k-th *signature tensor* by

$$\sigma^{(k)}(X) := \int_{0<t_1<\cdots<t_k<T} X'(t_1) \otimes \cdots \otimes X'(t_k)\ \mathrm{d}t_1 \cdots \mathrm{d}t_k \in (\mathbb{R}^d)^{\otimes k}$$

Example

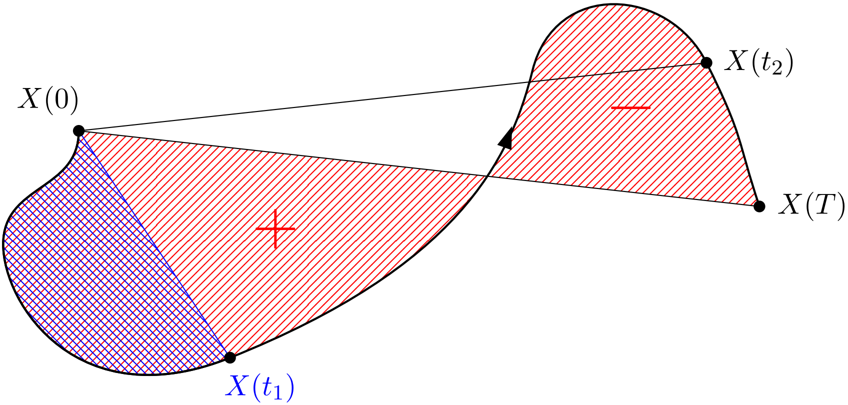
$$\sigma^{(1)}(X) = \int_0^T X'(t_1)\mathrm{d}t_1 = X(T) - X(0) \in \mathbb{R}^d$$

$$\begin{aligned}
\sigma^{(2)}(X) &= \int_{0<t_1<t_2<T} X'(t_1) \otimes X'(t_2)\, \mathrm{d}t_1\mathrm{d}t_2 \\
&= \int_0^T \int_0^{t_2} X'(t_1)\mathrm{d}t_1 \otimes X'(t_2)\mathrm{d}t_2 \\
&= \int_0^T (X(t_2) - X(0)) \otimes X'(t_2)\mathrm{d}t_2 \in \mathbb{R}^{d\times d}
\end{aligned}$$

Example

$$\begin{aligned}
&\operatorname{Sym}\sigma^{(2)}(X)\\
&=\frac{1}{2}\int_0^T (X(t_2)-X(0))\otimes X'(t_2)+X'(t_2)\otimes (X(t_2)-X(0))\mathrm{d}t_2\\
&=\frac{1}{2}(X(T)-X(0))\otimes (X(T)-X(0))
\end{aligned}$$

$$\begin{aligned}
&\operatorname{Skew}\sigma^{(2)}(X)\\
&=\frac{1}{2}\int_0^T (X(t_2)-X(0))\otimes X'(t_2)-X'(t_2)\otimes (X(t_2)-X(0))\mathrm{d}t_2\\
&=\,???
\end{aligned}$$



$$\text{SignedArea}(X^1, X^2)_t = \frac{1}{2}\Big(\int_0^t (X_1(s)-X_1(0))X_2'(s)-X_1'(s)(X_2(s)-X_2(0))\mathrm{d}$$

Picture from Diehl-Lyons-Preiß-Reizenstein, *Areas of areas generate the shuffle algebra*

$k \geqslant 3$

$$\sigma^{(k)}(X) = \operatorname{Sym}\sigma^{(k)}(X) + \operatorname{Skew}\sigma^{(k)}(X) + \text{ more}$$

$$\operatorname{Sym}\sigma^{(k)}(X) = \frac{1}{k!}(X(T) - X(0))^{\otimes k}$$

$$\operatorname{Skew}\sigma^{(k)}(X) =: \operatorname{SignedVolume}^{(k)}(X)$$

Why the fuzz?

We define the *iterated-integrals signature*

$$\sigma(X) := (1, \sigma^{(1)}(X), \sigma^{(2)}(X), \sigma^{(3)}(X), \dots).$$

Theorem (Chen)

If X is arclength-smooth, then $\sigma(X)$ uniquely characterizes X up to reparametrization and starting point among all arclenth-smooth paths.

Why the fuzz?

We define the *iterated-integrals signature*

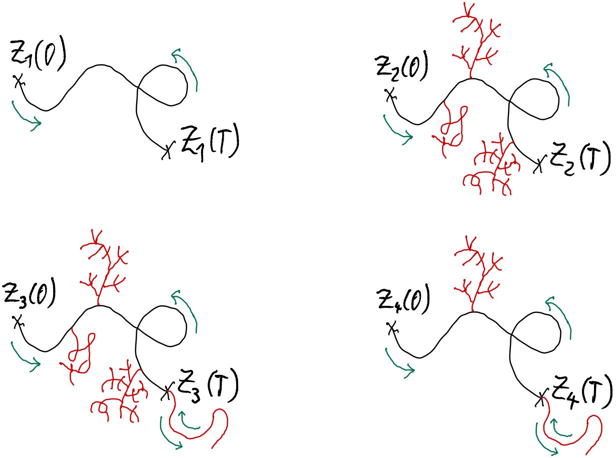
$$\sigma(X) := (1, \sigma^{(1)}(X), \sigma^{(2)}(X), \sigma^{(3)}(X), \dots).$$

Theorem (Chen)

If X is arclength-smooth, then $\sigma(X)$ uniquely characterizes X up to reparametrization and starting point among all arclenth-smooth paths.

What if X is not smooth?

Tree-like equivalence (aka thin homotopy equivalence)



We write $\check{X}$ for the reduced path of X.

We now consider the signature $\sigma(X)$ of a path $X : [0,T] \to \mathbb{R}^d$ as an element of $T((\mathbb{R}^d))$, the dual space of the tensor algebra $T(\mathbb{R}^d)$, i.e.

$$\langle \sigma(X), \mathtt{i}_1 \cdots \mathtt{i}_n \rangle = \int_{0<t_1<\cdots<t_n<T} X'_{i_1}(t_1) \cdots X'_{i_n}(t_n) \, \mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

We now consider the signature $\sigma(X)$ of a path $X : [0, T] \to \mathbb{R}^d$ as an element of $T((\mathbb{R}^d))$, the dual space of the tensor algebra $T(\mathbb{R}^d)$, i.e.

$$\langle \sigma(X), \mathtt{i}_1 \cdots \mathtt{i}_n \rangle = \int_{0<t_1<\cdots<t_n<T} X'_{i_1}(t_1) \cdots X'_{i_n}(t_n) \, \mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

Theorem (Chen's identity)

$$\sigma(X \sqcup Y) = \sigma(X) \bullet \sigma(Y).$$

Here, $\bullet$ is the internal tensor product of $T((\mathbb{R}^d))$, and $\sqcup$ is concatenation of paths.

Interpretation: σ is a semigroup homomorphism.

Many connections...

Algebraic Statistics $\rightsquigarrow$ Améndola, Friz, Sturmfels, *Varieties of Signature Tensors*

Machine Learning $\rightsquigarrow$ Chevyrev, Kormilitzin, *A Primer on the Signature Method in Machine Learning*

Toric geometry $\rightsquigarrow$ Colmenarejo, Galuppi, Michałek, *Toric geometry of path signature varieties*

Tropical geometry $\rightsquigarrow$ Diehl, Ebrahimi-Fard, Tapia, *Tropical time series, iterated-sums signatures and quasisymmetric functions*

Quantum physics $\rightsquigarrow$ Brown, *Iterated integrals in quantum field theory*

Previous work by algebraic geometers:
Améndola-Friz-Sturmfels et al: Study the complex projective Zariski closure of the finite dimensional semialgebraic set that is $\sigma^{(k)}(\mathcal{X}_\ell)$,

where $\mathcal{X}_\ell$ is piecewise linear paths/polynomial paths/log-linear rough paths of order ℓ.

Previous work by algebraic geometers:
Améndola-Friz-Sturmfels et al: Study the complex projective Zariski closure of the finite dimensional semialgebraic set that is $\sigma^{(k)}(\mathcal{X}_\ell)$,

where $\mathcal{X}_\ell$ is piecewise linear paths/polynomial paths/log-linear rough paths of order ℓ.

Our new approach:
Introduce a Zariski topology and algebraic geometry on the infinite dimensional path space itself.

In classical algebraic geometry, affine varieties in $\mathbb{R}^d$ are sets of the form

$$V(P) = \{x \in \mathbb{R}^d | p(x) = 0 \, \forall p \in P\}$$

where P is a set of polynomials $p : \mathbb{R}^d \to \mathbb{R}$.

Similarly, we now consider varieties in the space $\mathrm{BV}(\mathbb{R}^d)$ of continuous paths in $\mathbb{R}^d$ with bounded variation.

We call a *path variety* any subset of the form

$$\mathcal{V}(W) := \{X \in \mathrm{BV}(\mathbb{R}^d) | \langle \sigma(X), x \rangle = 0 \, \forall x \in W\}, \quad W \subseteq T(\mathbb{R}^d)$$

They form the closed sets of the *path Zariski topology*.

We turn $T(\mathbb{R}^d)$ into a free commutative associative algebra by introducing the shuffle product $\sqcup\!\sqcup$.
Let

$$\mathrm{Sh}(i,n) := \{\tau \in \mathrm{S}_n \,|\, \tau(1) < \ldots < \tau(i),\, \tau(i+1) < \ldots < \tau(n)\}$$

and

$$\mathtt{i}_1 \ldots \mathtt{i}_n \sqcup\!\sqcup \mathtt{i}_{n+1} \ldots \mathtt{i}_m := \sum_{\tau \in \mathrm{Sh}(n,m)} \mathtt{i}_{\tau^{-1}(1)} \cdots \mathtt{i}_{\tau^{-1}(m)}.$$

We turn $T(\mathbb{R}^d)$ into a free commutative associative algebra by introducing the shuffle product $\sqcup\!\sqcup$.
Let

$$\mathrm{Sh}(i,n) := \{\tau \in \mathrm{S}_n \,|\, \tau(1) < \ldots < \tau(i),\, \tau(i+1) < \ldots < \tau(n)\}$$

and

$$\mathtt{i}_1 \ldots \mathtt{i}_n \sqcup\!\sqcup\, \mathtt{i}_{n+1} \ldots \mathtt{i}_m := \sum_{\tau \in \mathrm{Sh}(n,m)} \mathtt{i}_{\tau^{-1}(1)} \cdots \mathtt{i}_{\tau^{-1}(m)}.$$

Example

$\mathtt{1} \sqcup\!\sqcup\, \mathtt{2} = \mathtt{12} + \mathtt{21}, \quad \mathtt{1} \sqcup\!\sqcup\, \mathtt{1} \sqcup\!\sqcup\, \mathtt{1} = 6 \cdot \mathtt{111}, \quad \mathtt{12} \sqcup\!\sqcup\, \mathtt{1} = \mathtt{121} + 2 \cdot \mathtt{112}$

Example

$$\begin{aligned}\langle \sigma(X), 1 ⧢ 2\rangle &= \langle \sigma(X), 12 + 21\rangle \\ &= \int_{0<t_1<t_2<T} X_1'(t_1)X_2'(t_2) + X_2'(t_1)X_1'(t_2)\,\mathrm{d}t_1\mathrm{d}t_2 \\ &= \int_0^T (X_1(t_2) - X_1(0))X_2'(t_2) + (X_2(t_2) - X_2(0)X_1'(t_2)\mathrm{d}t_2 \\ &= (X_1(T) - X_1(0))(X_2(T) - X_2(0)) \\ &= \langle \sigma(X), 1\rangle\langle \sigma(X), 2\rangle\end{aligned}$$

Example

$$\begin{aligned}\langle\sigma(X), 1 \sqcup\!\sqcup\, 2\rangle &= \langle\sigma(X), 12 + 21\rangle \\ &= \int_{0<t_1<t_2<T} X_1'(t_1)X_2'(t_2) + X_2'(t_1)X_1'(t_2)\,\mathrm{d}t_1\mathrm{d}t_2 \\ &= \int_0^T (X_1(t_2) - X_1(0))X_2'(t_2) + (X_2(t_2) - X_2(0)X_1'(t_2)\mathrm{d}t_2 \\ &= (X_1(T) - X_1(0))(X_2(T) - X_2(0)) \\ &= \langle\sigma(X), 1\rangle\langle\sigma(X), 2\rangle\end{aligned}$$

Theorem (Ree's shuffle identity)

$$\langle\sigma(X), x \sqcup\!\sqcup\, y\rangle = \langle\sigma(X), x\rangle\langle\sigma(X), y\rangle$$

We call a *path variety* any subset of the form

$$\mathcal{V}(W) := \{X \in \mathrm{BV}(\mathbb{R}^d) | \langle \sigma(X), x \rangle = 0 \, \forall x \in W\}, \quad W \subseteq T(\mathbb{R}^d)$$

They form the closed sets of the *path Zariski topology*.

We call a *path variety* any subset of the form

$$\mathcal{V}(W) := \{X \in \mathrm{BV}(\mathbb{R}^d) | \langle \sigma(X), x \rangle = 0 \, \forall x \in W\}, \quad W \subseteq T(\mathbb{R}^d)$$

They form the closed sets of the *path Zariski topology*.

Path varieties are in 1-to-1 correspondance to the BV 'radical' shuffle ideals

$$\mathcal{I}(U) := \{x \in T(\mathbb{R}^d) | \langle \sigma(X), x \rangle = 0 \, \forall X \in U\}, \quad U \subseteq \mathrm{BV}(\mathbb{R}^d).$$

$\mathcal{V} \circ \mathcal{I}$ is the closure operator, and $\mathcal{I} \circ \mathcal{V}$ is the BV radical operator.

Our full algebraic and combinatorial structure is

$$(T(\mathbb{R}^d), \sqcup\!\sqcup, \Delta_\bullet, \mathcal{A}, \succ, \prec).$$

Let the right $\succ$ and left $\prec$ halfshuffles be recursively defined by

$$w \succ \mathtt{i} := w\mathtt{i}, \qquad \mathtt{i} \prec w := \mathtt{i}w$$
$$w \succ v\mathtt{i} := (w \succ v + v \succ w)\mathtt{i}, \qquad \mathtt{i}v \prec w := \mathtt{i}(w \prec v + v \prec w)$$

Then

$$x \sqcup\!\sqcup y = x \succ y + y \succ x = x \prec y + y \prec x$$

and

$$\mathcal{A}(x \succ y) = \mathcal{A}y \prec \mathcal{A}x, \quad \mathcal{A}(x \prec y) = \mathcal{A}y \succ \mathcal{A}x.$$

Let $\langle W \rangle_\succ$ denote the two-sided $\succ$-ideal generated by W.

Example

$$
\begin{aligned}
&\langle \sigma(X), w \succ 1\rangle = \langle \sigma(X), w1\rangle \\
&= \int_{0<t_1<\cdots<t_n<t<T} X'_{i_1}(t_1)\ldots X'_{i_n}(t_n)X_1(t)\,\mathrm{d}t_1\ldots\mathrm{d}t_n\mathrm{d}t \\
&= \int_0^T \int_{0<t_1<\cdots<t_n<t} X'_{i_1}(t_1)\ldots X'_{i_n}(t_n)\,\mathrm{d}t_1\ldots\mathrm{d}t_n X_1(t)\mathrm{d}t \\
&= \int_0^T \langle \sigma(X)_t, w\rangle\langle \sigma(X)'_t, 1\rangle\,\mathrm{d}t
\end{aligned}
$$

Example

$$\begin{aligned}
&\langle \sigma(X), w \succ 1 \rangle = \langle \sigma(X), w1 \rangle \\
&= \int_{0<t_1<\cdots<t_n<t<T} X'_{i_1}(t_1)\ldots X'_{i_n}(t_n) X_1(t)\,\mathrm{d}t_1\ldots\mathrm{d}t_n\mathrm{d}t \\
&= \int_0^T \int_{0<t_1<\cdots<t_n<t} X'_{i_1}(t_1)\ldots X'_{i_n}(t_n)\,\mathrm{d}t_1\ldots\mathrm{d}t_n X_1(t)\mathrm{d}t \\
&= \int_0^T \langle \sigma(X)_t, w\rangle \langle \sigma(X)'_t, 1\rangle\,\mathrm{d}t
\end{aligned}$$

In general,

$$\langle \sigma(X), x \succ y \rangle = \int_0^T \langle \sigma(X)_t, x\rangle \langle \sigma(X)'_t, y\rangle\,\mathrm{d}t$$

Example

$$\begin{aligned}&\langle\sigma(X), w \succ 1\rangle = \langle\sigma(X), w1\rangle \\ &= \int_{0<t_1<\cdots<t_n<t<T} X'_{i_1}(t_1)\ldots X'_{i_n}(t_n)X_1(t)\,\mathrm{d}t_1\ldots\mathrm{d}t_n\mathrm{d}t \\ &= \int_0^T\int_{0<t_1<\cdots<t_n<t} X'_{i_1}(t_1)\ldots X'_{i_n}(t_n)\,\mathrm{d}t_1\ldots\mathrm{d}t_n X_1(t)\mathrm{d}t \\ &= \int_0^T \langle\sigma(X)_t, w\rangle\langle\sigma(X)'_t, 1\rangle\,\mathrm{d}t\end{aligned}$$

In general,

$$\langle\sigma(X), x \succ y\rangle = \int_0^T \langle\sigma(X)_t, x\rangle\langle\sigma(X)'_t, y\rangle\,\mathrm{d}t$$

So $\succ$ is an abstraction of $(f, g) \mapsto \int_0^{\cdot} f(t)g'(t)\mathrm{d}t$

Theorem (Preiß)

Whenever a set of paths U contains history, i.e. all left subpaths of reduced paths, $\mathcal{I}(U)$ is a $\succ$-ideal.

Whenever I is a $\succ$-ideal, $\mathcal{V}(I)$ contains history.

Corollary

Let $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with $p(0) = 0$. Then $\mathcal{V}(\langle p_i^{\sqcup\!\sqcup} \rangle_{\succ})$ is the variety of all paths X such that $\check{X} - X_0$ lies in the point variety M defined by the vanishing of all p_i.

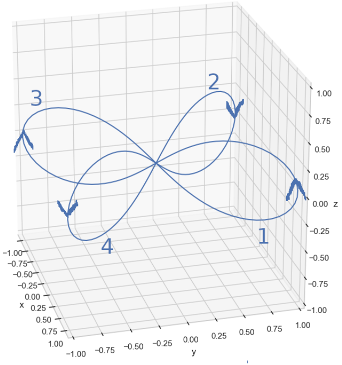
Example

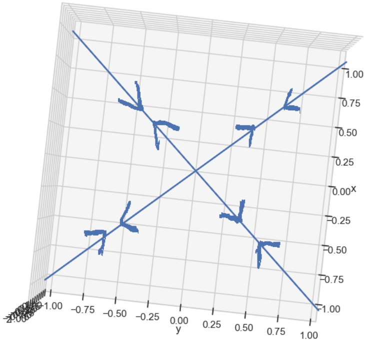
The variety of all paths starting in $(0, 0)$ and staying on the unit circle centered in $(1, 0)$ is described by the halfshuffle ideal

$$\langle \mathtt{11} - \mathtt{1} + \mathtt{22} \rangle_{\succ}$$

Example

While any path (X^1, X^2, X^3) which is tree-like equivalent to $(0, 0, X^3)$ and $(X^1, 0, 0)$ is tree-like, there are non-tree-like paths (X^1, X^2, X^3) such that (X^1, X^2), (X^1, X^3) and (X^2, X^3) are all tree-like.





Theorem (Preiß)

If $M \subseteq \mathrm{BV}(\mathbb{R}^d)$ is a set of paths closed under concatenation, then the variety $\bar{M}$ is closed under concatenation, time reversal and taking admissible roots, and $\mathcal{I}(M)$ is a Hopf ideal.

Corollary

The set of lattice paths $\mathfrak{L}$ is Zariski dense in $\mathrm{BV}(\mathbb{R}^d)$.

V contains history
$\Rightarrow \mathcal{I}(V)$ is a halfshuffle ideal
$\Rightarrow \mathbb{R}[V] := T(\mathbb{R}^d)/\mathcal{I}(V)$ is a halfshuffle algebra

V is stable under concatenation
$\Rightarrow \mathcal{I}(V)$ is a Hopf ideal
$\Rightarrow \mathbb{R}[V] := T(\mathbb{R}^d)/\mathcal{I}(V)$ is a Hopf algebra

Note: To understand the geometrical structure of V, we need the algebraic structure of $\mathbb{R}[V]$ **plus** the BV radical operator on the power set of $\mathbb{R}[V]$.

We can do polynomial ODEs!

$f' = f + 1, f(0) = 0 \rightsquigarrow f(t) = \int_0^t f(t) + 1 \mathrm{d}t \rightsquigarrow \left\langle \mathtt{1} - \mathtt{10} - \mathtt{0} \right\rangle_{\succ}$
with solution space $\{[0,T] \to \mathbb{R} | t \mapsto (t, \exp(t) - 1)\}$

$f'' = -f - 1, f(0) = 0, f'(0) = 0 \rightsquigarrow f(t) = \int_0^t \int_0^s (-f(t) - 1) \mathrm{d}s \mathrm{d}t$
$\rightsquigarrow \left\langle \mathtt{1} + \mathtt{100} + \mathtt{00} \right\rangle_{\succ}$
with solution space $\{[0,T] \to \mathbb{R} | t \mapsto (t, \cos(t) - 1)\}$

We can do polynomial ODEs!

$f' = f + 1, f(0) = 0 \rightsquigarrow f(t) = \int_0^t f(t) + 1\mathrm{d}t \rightsquigarrow \left\langle \mathtt{1} - \mathtt{10} - \mathtt{0} \right\rangle_{\succ}$
with solution space $\{[0,T] \to \mathbb{R} | t \mapsto (t, \exp(t) - 1)\}$

$f'' = -f - 1, f(0) = 0, f'(0) = 0 \rightsquigarrow f(t) = \int_0^t \int_0^s (-f(t) - 1)\mathrm{d}s\mathrm{d}t$
$\rightsquigarrow \left\langle \mathtt{1} + \mathtt{100} + \mathtt{00} \right\rangle_{\succ}$
with solution space $\{[0,T] \to \mathbb{R} | t \mapsto (t, \cos(t) - 1)\}$

$f' = f^2 + 1, f(0) = 0 \rightsquigarrow f(t) = \int_0^t f^2(t) + 1\mathrm{d}t \rightsquigarrow \left\langle \mathtt{1} - 2 \cdot \mathtt{110} - \mathtt{0} \right\rangle_{\succ}$
with solution space $\{[0,T] \to \mathbb{R} | t \mapsto (t, \tan(t))\}$ (explosion!)

$f' = \sqrt{f}, f(0) = 0 \rightsquigarrow f(t) = \int_0^t g(t)\mathrm{d}t, f = g^2 \rightsquigarrow \left\langle \mathtt{1} - \mathtt{20}, \mathtt{1} - 2 \cdot \mathtt{22} \right\rangle_{\succ}$
with solution space $\{[0,T] \to \mathbb{R} | t \mapsto (t, 1_{t \geqslant c} \cdot \frac{t^2}{4}, 1_{t \geqslant c} \cdot \frac{t}{2})\}$
(non-uniqueness!)

More to come...

- abstract path varieties
- morphisms between affine/abstract path varieties
- complex and projective path varieties
- rough paths on point varieties
- generalized point varieties
- semi-algebraic path sets
- study of singularities
- algebraic path groupoids

Thank you

Check out my website: rosapreiss.net

And the paper:

An algebraic geometry of paths via the iterated-integrals signature
arXiv:2311.17886[math.RA]