Introduction to signatures

Nikolas Tapia

NTNU Trondheim

Feb. 26, 2019 @ Magic 2019, Ilsetra

Goals

1. Signatures
    1. For paths on $\mathbb{R}^d$
    2. For paths on a Lie group
    3. Practical questions
2. Rough paths
    1. Geometric rough paths
    2. Branched rough paths
    3. Problem of existence

Signatures

Signatures were studied by K.-T. Chen in the 50's in order to clasify smooth curves on manifolds.

They were later generalized by T. Lyons by the end of the 90's to what he called *rough paths*.

Almost 20 years later, they have found many applications in Machine Learning, Data Analysis and trend recognition in time series.

Consider a d-dimensional vector space V and define

$$T(V) := \mathbb{R}\mathbf{1} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots .$$

For $p \geq 1$, the *degree p component* $T(V)_p = V^{\otimes p}$ is spanned by the set

$$\{e_{i_1 \cdots i_p} := e_{i_1} \otimes \cdots \otimes e_{i_p} : i_1, \ldots, i_p = 1, \ldots, d\}$$

In particular $\dim T(V) = \infty$.

For a given $\psi \in T(V)^* := T(\!(V)\!)$ we write

$$\psi = \sum_{p \geq 0} \sum_{i_1, \ldots, i_p = 1}^{d} \langle \psi, e_{i_1 \cdots i_p} \rangle e_{i_1 \cdots i_p} .$$

There are two products on $T(V)$:

1. the tensor product: $e_{i_1\cdots i_p}\otimes e_{i_{p+1}\cdots i_{p+q}}=e_{i_1\cdots i_{p+q}}\in T(V)_{p+q}$ and,
2. the shuffle product:

$$e_{i_1\cdots i_p}\sqcup\!\sqcup\, e_{i_{p+1}\cdots i_{p+q}}=\sum_{\sigma\in\mathrm{Sh}(p,q)}e_{i_{\sigma(1)}i_{\sigma(2)}\cdots i_{\sigma(p+q)}}\in T(V)_{p+q}.$$

Examples:

$$e_i\sqcup\!\sqcup\, e_j=e_{ij}+e_{ji},\quad e_i\sqcup\!\sqcup\, e_{jk}=e_{ijk}+e_{jik}+e_{jki}.$$

On both cases $1\in T(V)$ acts as the unit.

The shuffle algebra carries a coalgebra structure: define $\Delta : T(V) \to T(V) \dot{\otimes} T(V)$ by

$$\Delta e_{i_1 \cdots i_p} := e_{i_1 \cdots i_p} \dot{\otimes} 1 + 1 \dot{\otimes} e_{i_1 \cdots i_p} + \sum_{j=1}^{p-1} e_{i_1 \cdots i_j} \dot{\otimes} e_{i_{j+1} \cdots i_p}.$$

This structure is dual to the tensor product in the sense that if $\varphi, \psi \in T(\!(V)\!)$ then

$$\varphi \otimes \psi = \sum_{p \geq 0} \sum_{i_1, \ldots, i_p = 1}^{d} \langle \varphi \dot{\otimes} \psi, \Delta e_{i_1 \cdots i_p} \rangle e_{i_1 \cdots i_p}.$$

Let $x : [0,1] \to \mathbb{R}^d$ be a curve with finite 1-variation.

Its signature over the interval $[s,t]$ is the tensor series with coefficients

$$\langle S(x)_{s,u}, 1\rangle = 1, \quad \langle S(x)_{s,t}, e_{i_1\cdots i_p} \otimes e_j\rangle := \int_s^t \langle S(x)_{s,u}, e_{i_1\cdots i_p}\rangle \dot{x}_u^j \,\mathrm{d}u.$$

Example:

$$\langle S(x)_{s,t}, e_i\rangle = \int_s^t \dot{x}_u^i \,\mathrm{d}u = x_t^i - x_s^i, \quad \langle S(x)_{s,t}, e_{ij}\rangle = \int_s^t \int_s^u \dot{x}_v^i \dot{x}_u^j \mathrm{d}v \mathrm{d}u.$$

Chen shows that $S(x)$ satsifies:

1. the shuffle relation:

$$\langle S(x)_{s,t}, e_{i_1\ldots i_p} \sqcup\!\sqcup\, e_{i_{p+1}\cdots i_{p+q}} \rangle = \langle S(x)_{s,t}, e_{i_1\ldots i_p} \rangle \langle S(x)_{s,t}, e_{i_{p+1}\cdots i_{p+q}} \rangle.$$

2. Chen's rule: if y is another path and $x \cdot y$ is their concatenation, then

$$S(x \cdot y)_{s,t} = S(x)_{s,u} \otimes S(y)_{u,t}.$$

Moreover, one can show that there exists a constant $C > 0$ such that

$$|\langle S(x)_{s,t}, e_{i_1\cdots i_p} \rangle| \leq \frac{C^p}{p!} |t-s|^p.$$

For example

$$\begin{aligned}
\langle S(x)_{s,t}, e_{ij} + e_{ji} \rangle &= \int_s^t \int_s^u \dot{x}_v^i \dot{x}_u^j \, \mathrm{d}v \mathrm{d}u + \int_s^t \int_s^u \dot{x}_v^j \dot{x}_u^i \, \mathrm{d}v \mathrm{d}u \\
&= \int_s^t (x_u^i - x_s^i) \dot{x}_u^j \, \mathrm{d}u + \int_s^t (x_u^j - x_s^j) \dot{x}_u^i \, \mathrm{d}u \\
&= (x_t^i - x_s^i)(x_t^j - x_s^j) \\
&= \langle S(x)_{s,t}, e_i \rangle \langle S(x)_{s,t}, e_j \rangle.
\end{aligned}$$

Another example:

$$
\begin{aligned}
\langle S(x)_{s,t}, e_i \rangle &= \int_s^t \dot{x}_v^i \, \mathrm{d}v \\
&= \int_s^u \dot{x}_v^i \, \mathrm{d}v + \int_u^t \dot{x}_v^i \, \mathrm{d}v \\
&= \langle S(x)_{s,u} \otimes S(x)_{u,t}, e_i \rangle.
\end{aligned}
$$

Yet another example:

$$\begin{aligned}
\langle S(x)_{s,t}, e_{ij} \rangle &= \int_s^t (x_v^i - x_s^i)\dot{x}_v^j \,\mathrm{d}v \\
&= \int_s^u (x_v^i - x_s^i)\dot{x}_v^j \,\mathrm{d}v + \int_u^t (x_v^i - x_s^i)\dot{x}_v^j \,\mathrm{d}v \\
&= \int_s^u (x_v^i - x_s^i)\dot{x}_v^j \,\mathrm{d}v + \int_u^t (x_v^i - x_u^i)\dot{x}_v^j \,\mathrm{d}v + (x_u^i - x_s^i)(x_t^j - x_u^j) \\
&= \langle S(x)_{s,u} \otimes S(x)_{u,t}, e_{ij} \rangle
\end{aligned}$$

Signatures can be easiy computed for certain paths.

If x is a straight line, i.e. $x_t = a + bt$ with $a, b \in \mathbb{R}^d$ then

$$\langle S(x)_{s,t}, e_{i_1 \cdots i_p} \rangle = \frac{(t-s)^p}{p!} \prod_{j=1}^{p} b_{i_j}.$$

Indeed

$$\begin{aligned}\langle S(x)_{s,t}, e_{i_1 \cdots i_p} \otimes e_j \rangle &= \int_s^t \frac{(u-s)^p}{p!} \prod_{k=1}^{p} b_{i_k} b_j \, \mathrm{d}u \\ &= \frac{(t-s)^{p+1}}{(p+1)!} \prod_{k=1}^{p+1} b_{i_k}.\end{aligned}$$

Therefore

$$S(x)_{s,t} = 1 + (t-s)b + \frac{(t-s)^2}{2} b \otimes b + \frac{(t-s)^3}{6} b \otimes b \otimes b + \cdots = \exp_\otimes((t-s)b).$$

By Chen's rule, if x is a general piecewise linear path with slopes $b_1, \ldots, b_m \in \mathbb{R}^d$ between times $s < t_1 < \cdots < t_{m-1} < t$ then

$$S(x)_{s,t} = \exp_\otimes((t_1 - s)b_1) \otimes \cdots \otimes \exp_\otimes((t - t_{m-1})b_m).$$

Some further properties:

1. Invariant under reparametrization: if φ is an increasing diffeomorphism on $[0,1]$ then
$$S(x \circ \varphi)_{s,t} = S(x)_{s,t}.$$
2. Characterizes the path up-to *irreducibility*. If $S(x) = S(y)$ for two irreducible paths then y is a translation of x.

For some applications one has a set of points sampled in time.

One may construct the signature by linear interpolation.

How to deal with infinite series? Truncation is one possibility, but to which level?

Applying some transformations one can read off some information from the signature:

1. Mean
2. Quadratic variation, i.e. variance

For Machine Learning applications, levels of the signature are selected as explanatory variables for the features of a path. An example of objective function (taken from Gyurkó, Lyons, Kontkowski & Field; 2014)

$$\min_{\beta}\left[\sum_{k=1}^{L}\left(\sum_{|w|\leq M}\beta_w\langle S(x_k)_{0,1},w\rangle - y_k\right)^2 + \alpha\sum_{|w|\leq M}|\beta_w|\right]$$

Let G be a d-dimensional Lie group with Lie algebra $\mathfrak{g}$.

The Maurer–Cartan form on G is the pushforward of left translation:

$$\omega_g(v) = (L_{g^{-1}})_* v, \quad v \in T_g G.$$

It is a $\mathfrak{g}$-valued 1-form, i.e. a smooth section of $(M \times \mathfrak{g}) \otimes T^*G$. In other words, ω_g maps $T_g G$ into $\mathfrak{g}$. In particular, it can be written as

$$\omega = X_1 \otimes \omega^1 + \cdots + X_d \otimes \omega^d$$

where $\omega^1, \ldots, \omega^d$ are suitable 1-forms on G and $X_1, \ldots, X_d$ is a basis of $\mathfrak{g}$.

Chen defines the signature over the interval $[s,t]$ of a smooth curve $\alpha:[0,1]\to G$ as the tensor series $S(\alpha)_{s,t}$ with coefficients

$$\langle S(\alpha)_{s,t},1\rangle := 1,\quad \langle S(\alpha)_{s,t},e_{i_1\dots i_p}\otimes e_j\rangle := \int_s^t \langle S(\alpha)_{s,u},e_{i_1\cdots i_p}\rangle \omega^j_{\alpha_u}(\dot\alpha_u)\,\mathrm{d}u.$$

When $G=\mathbb{R}^d$ this definition coincides with the previous one by observing that $\omega^i=\mathrm{d}x^i$, i.e.

$$\omega_{\alpha_t}(\dot\alpha_t)=\dot\alpha_t^1 e_1+\cdots+\dot\alpha_t^d e_d$$

with $e_1,\dots,e_d$ the canonical basis of $\mathbb{R}^d$.

An example: let $G = H_3$ be the Heisenberg group, that is,

$$H_3 := \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{R} \right\}.$$

Its Lie algebra $\mathfrak{h}_3$ is spanned by the matrices

$$X := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with $[X, Y] = Z$, $[X, Z] = [Y, Z] = 0$.

In this group, the Maurer–Cartan form is given by

$$\omega_g = \begin{pmatrix} 0 & \mathrm{d}x & \mathrm{d}z - x\mathrm{d}y \\ 0 & 0 & \mathrm{d}y \\ 0 & 0 & 0 \end{pmatrix} \text{ when } g = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

In particular

$$S(\alpha)_{s,t} = 1 + \int_s^t \dot{\alpha}_u^x \, \mathrm{d}u \; e_1 + \int_s^t \dot{\alpha}_u^y \, \mathrm{d}u \; e_2 + \int_s^t (\dot{\alpha}_u^z - \alpha_u^x \dot{\alpha}_u^y) \, \mathrm{d}u \; e_3 + \cdots \in T(\!(\mathbb{R}^3)\!)$$

where

$$\alpha_t = \begin{pmatrix} 1 & \alpha_t^x & \alpha_t^z \\ 0 & 1 & \alpha_t^y \\ 0 & 0 & 1 \end{pmatrix}.$$

Rough paths

Definition (Lyons, 1998)

Let $p > 1$. A p-geometric rough path is a collection $(X_{s,t} : s, t \in [0, 1])$ of linear maps on the shuffle algebra satisfying Chen's rule, the shuffle identity and the analitic bound

$$|\langle X_{s,t}, e_{i_1 \cdots i_p} \rangle| \leq C_p |t - s|^{\gamma/p}.$$

We say X is a GRP above a path $x : [0, 1] \to \mathbb{R}^d$ if $\langle X_{s,t}, e_i \rangle = x_t^i - x_s^i$.

Problem: Existence is not obvious. Can be constructed on specific cases, for example Brownian Motion, fractional Brownian Motion with Hurst parameter $\frac{1}{3} < H < \frac{1}{2}$, etc.

Let $N := \lceil p \rceil$.

Theorem (Lyons, 1998)

Let X be a partial homomorphism defined only on $T(V)_1 \oplus \cdots \oplus T(V)_N$, satisfying the analitical constraint and Chen's rule. There exists a unique extension $\hat{X}$ to all of $T(V)$ satisfying the definition of a geometric rough path.

Theorem (Lyons–Victoir, 2007)

Given a path of finite p-variation, there exists a p-geometric rough path above x.

Geometric rough paths provide a "universal" description of flows controlled by x.

For a 1-dimensional smooth path x, consider the controlled differential equation

$$\dot{y}_t = V(y_t)\dot{x}_t.$$

To first order we have

$$y_t - y_s = V(y_s)\int_s^t \dot{x}_u \,\mathrm{d}u + o(|t-s|)$$

To second order

$$y_t - y_s = V(y_s)\int_s^t \dot{x}_u \,\mathrm{d}u + V'(y_s)V(y_s)\int_s^t\int_s^u \dot{x}_v\dot{x}_u \,\mathrm{d}v\mathrm{d}u + o(|t-s|^2).$$

Lyons' main goal was to treat Stochastic Differential Equations.

We know that Brownian motion a.s. has finite p-variation for any $p > 2$.

Thus, if we fix $2 < p < 3$, then for any *fixed* realization

$$X_{s,t} = 1 + \sum_{i=1}^{d} (B_t^i - B_s^i)\ e_i + \sum_{i,j=1}^{d} \int_s^t (B_u^i - B_s^i) \circ \mathrm{d}B_u^j\ e_{ij}$$

satisfies the required hypothesis.

Therefore we have a notion of path-wise solution to an SDE via rough paths.

The shuffle relations forbid considering Itô-type SDEs.

The correct framework for this is **branched rough paths** introduce by Gubinelli in 2010.

Again, the problem of existence arises:

Theorem (T.–Zambotti, 2018)

Given a γ-Hölder path $x : [0,1] \to \mathbb{R}^d$, there exists a branched rough path above x.

Moreover

Theorem

There is a Lie group $\mathscr{C}^\gamma$ acting freely and transitively on the space $\mathbf{BRP}^\gamma$ of branched rough paths. In particular, $\mathbf{BRP}^\gamma$ is a principal homogeneous space for $\mathscr{C}^\gamma$.