Signatures methods in finance

Christa Cuchiero
partly based on a course given jointly with Sara Svaluto-Ferro

University of Vienna

Mini course

Soesterberg, January 2024

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference?

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Evidence = Data

- Time series data
- Derivatives' price data
- Macro economic data
- ...

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Evidence = Data

- Time series data
- Derivatives' price data
- Macro economic data
- ...

Reasoning

- Recognition of universal structures (statistics)
- First principles, e.g. no arbitrage
- Universal model classes and strategies

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Evidence = Data

- Time series data
- Derivatives' price data
- Macro economic data
- ...

Reasoning

- Recognition of universal structures (statistics)
- First principles, e.g. no arbitrage
- Universal model classes and strategies

Combining machine learning with theory from mathematical finace allows to conciliate both sides - modeling as close as possible to high dimensional data while obeying well established principles.

Machine learning ingredients

What are the ingredients to achieve this?

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.

Machine learning ingredients

What are the ingredients to achieve this?

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.

Machine learning ingredients

What are the ingredients to achieve this?

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.
2. Optimization criterion coming with a
   - loss function taylored to the problem.

Machine learning ingredients

What are the ingredients to achieve this?

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.
2. Optimization criterion coming with a
   - loss function taylored to the problem.
3. Algorithm used for training, typically
   - (stochastic) gradient type algorithms;
   - linear regression methods (if the regression basis is linear);
   - tools from convex (quadratic) optimization (if the problem allows for such a formulation).

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems; control problems in finance comprise portfolio optimization, hedging, optimal execution, optimal stopping, etc.

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems; control problems in finance comprise portfolio optimization, hedging, optimal execution, optimal stopping, etc.
  2. universal model classes of dynamic processes, being able to approximate classical financial models.

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems; control problems in finance comprise portfolio optimization, hedging, optimal execution, optimal stopping, etc.
  2. universal model classes of dynamic processes, being able to approximate classical financial models.
- The optimization criteria and loss functions depend on the problem and include
  - maximizing expected utility;
  - minimizing a risk measure;
  - maximizing over stopping times e.g. for pricing American options;
  - minimizing certain distances to time series and option price data $\Rightarrow$ calibration functionals.

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems; control problems in finance comprise portfolio optimization, hedging, optimal execution, optimal stopping, etc.
  2. universal model classes of dynamic processes, being able to approximate classical financial models.
- The optimization criteria and loss functions depend on the problem and include
  - maximizing expected utility;
  - minimizing a risk measure;
  - maximizing over stopping times e.g. for pricing American options;
  - minimizing certain distances to time series and option price data $\Rightarrow$ calibration functionals.
- As the regression basis is linear, many problems reduce to linear regression or convex quadratic optimization problems.

Themes of this lecture

Part I Introduction to the theory of signature

- Signature of continuous bounded variation paths and the Lie group structures;
- Continuous rough paths, continuous semimartingales as continuous rough paths and their signature;
- Universal approximation property of linear functions of the signature in appropriate topologies on path space.

Themes of this lecture

Part I Introduction to the theory of signature

- Signature of continuous bounded variation paths and the Lie group structures;
- Continuous rough paths, continuous semimartingales as continuous rough paths and their signature;
- Universal approximation property of linear functions of the signature in appropriate topologies on path space.

Part II Signature methods in stochastic portfolio theory

- Signature-type portfolios,
- Optimization tasks and approximation results;
- Numerical results on simulated and real market data.

Themes of this lecture

Part III An affine and polynomial perspective to signature based models

- An introduction to affine and polynomial processes
- Signature based models as affine and polynomial processes

Themes of this lecture

Part III An affine and polynomial perspective to signature based models

- An introduction to affine and polynomial processes
- Signature based models as affine and polynomial processes

Part IV Signature based asset price models for the joint calibration problem to SPX and VIX options

- Theoretical model ingredients and pricing of VIX and SPX options
- Calibration results on market data

Themes of this lecture

Part III An affine and polynomial perspective to signature based models

- An introduction to affine and polynomial processes
- Signature based models as affine and polynomial processes

Part IV Signature based asset price models for the joint calibration problem to SPX and VIX options

- Theoretical model ingredients and pricing of VIX and SPX options
- Calibration results on market data

Part V Signature of càdlàg rough paths and Lévy type signature models

- Extension of universal approximation results to càdlàg rough paths;
- Lévy driven signature based asset price models;
- Signature jump diffusions as affine and polynomial processes.

Part I

Introduction to the theory of signature

partly based on Chapter 7 of "Multidimensional stochastic processes as rough paths - Theory and Applications" by Friz & Victoir (2010)

Continuous paths of bounded variation

- Let $T > 0$ and let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_k = T\}$ denote a partition of $[0, T]$ and $\sum_{t_i \in \mathcal{D}}$ the summation over all points in $\mathcal{D}$.

Continuous paths of bounded variation

- Let $T > 0$ and let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_k = T\}$ denote a partition of $[0, T]$ and $\sum_{t_i \in \mathcal{D}}$ the summation over all points in $\mathcal{D}$.
- We define the 1-variation of a path $X \in C([0, T], \mathbb{R}^d)$ by

$$\|X\|_{1-var} := \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_i \in \mathcal{D}} |X_{t_i} - X_{t_{i+1}}|_{\mathbb{R}^d} \right),$$

  where $|\cdot|_{\mathbb{R}^d}$ denotes the Euclidian distance, i.e. $|a|_{\mathbb{R}^d} = \sqrt{\sum_{i=1}^d (a^i)^2}$ for $a \in \mathbb{R}^d$.
- We often write $X_{s,t}$ for the increment $X_t - X_s$, so that we can also write $\|X\|_{1-var} = \sup_{\mathcal{D} \subset [0,T]} (\sum_{t_i \in \mathcal{D}} |X_{t_i, t_{i+1}}|_{\mathbb{R}^d})$.

Continuous paths of bounded variation

- Let $T > 0$ and let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_k = T\}$ denote a partition of $[0, T]$ and $\sum_{t_i \in \mathcal{D}}$ the summation over all points in $\mathcal{D}$.
- We define the 1-variation of a path $X \in C([0, T], \mathbb{R}^d)$ by

$$\|X\|_{1-var} := \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_i \in \mathcal{D}} |X_{t_i} - X_{t_{i+1}}|_{\mathbb{R}^d} \right),$$

  where $|\cdot|_{\mathbb{R}^d}$ denotes the Euclidian distance, i.e. $|a|_{\mathbb{R}^d} = \sqrt{\sum_{i=1}^d (a^i)^2}$ for $a \in \mathbb{R}^d$.
- We often write $X_{s,t}$ for the increment $X_t - X_s$, so that we can also write $\|X\|_{1-var} = \sup_{\mathcal{D} \subset [0,T]} (\sum_{t_i \in \mathcal{D}} |X_{t_i, t_{i+1}}|_{\mathbb{R}^d})$.
- If $\|X\|_{1-var} < \infty$, we say that X is of bounded variation or of finite (1-) variation on $[0, T]$.
- The space of continuous paths of finite (1)-variation on $[0, T]$ with values in $\mathbb{R}^d$ is denoted by $C^{1-var}([0, T], \mathbb{R}^d)$.

Iterated integrals for continuous bounded variation paths

Definition

Let $X : [0, T] \to \mathbb{R}^d$ be a continuous path of bounded variation. Moreover, for $n \in \mathbb{N}$ let $I = (i_1, \ldots, i_n)$ be a multi-index with entries in $\{1, \ldots, d\}$.

Iterated integrals for continuous bounded variation paths

Definition

Let $X : [0, T] \to \mathbb{R}^d$ be a continuous path of bounded variation. Moreover, for $n \in \mathbb{N}$ let $I = (i_1, \ldots, i_n)$ be a multi-index with entries in $\{1, \ldots, d\}$. For $0 \leq s < t \leq T$, define the n-th iterated integrals of the path segment $X_{[s,t]}$ as

$$\mathbb{X}^{(n)}_{s,t;I} := \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dX^{i_1}_{u_1} \cdots dX^{i_n}_{u_n} \in \mathbb{R}$$

and denote by $\mathbb{X}^{(n)}_{s,t} \in (\mathbb{R}^d)^{\otimes n} \cong \mathbb{R}^{d^n}$ the collection of such integrals for multi-indices of length n.

Iterated integrals for continuous bounded variation paths

Definition

Let $X : [0, T] \to \mathbb{R}^d$ be a continuous path of bounded variation. Moreover, for $n \in \mathbb{N}$ let $I = (i_1, \dots, i_n)$ be a multi-index with entries in $\{1, \dots, d\}$. For $0 \leq s < t \leq T$, define the n-th iterated integrals of the path segment $X_{[s,t]}$ as

$$\mathbb{X}^{(n)}_{s,t;I} := \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dX^{i_1}_{u_1} \cdots dX^{i_n}_{u_n} \in \mathbb{R}$$

and denote by $\mathbb{X}^{(n)}_{s,t} \in (\mathbb{R}^d)^{\otimes n} \cong \mathbb{R}^{d^n}$ the collection of such integrals for multi-indices of length n. Then

$$\mathbb{X}_{s,t} = (1, \mathbb{X}^{(1)}_{s,t}, \dots, \mathbb{X}^{(n)}_{s,t}, \dots)$$

is called signature of the path segment $X_{[s,t]}$ and $\mathbb{X}^N_{s,t}$ denotes the step-N signature, i.e. the signature truncated at level N.

Understanding the iterated integrals - simplest example

- Consider the one-dimensional situation $d = 1$. In this case $(\mathbb{R}^1)^{\otimes n} \cong \mathbb{R}$.
- We now compute the signature of the path $t \mapsto X_t = t$.
- Then we obtain via induction

$$
\begin{aligned}
\mathbb{X}_{s,t}^{(1)} &= \int_s^t du = t - s \\
\mathbb{X}_{s,t}^{(2)} &= \int_s^t \int_s^{u_2} du_1 du_2 = \int_s^t (u_2 - s) du_2 = \frac{(t-s)^2}{2}. \\
&\vdots \\
\mathbb{X}_{s,t}^{(n)} &= \int_s^t \Big(\int_s^{u_n} \cdots \int_s^{u_2} du_1 \cdots du_{n-1} \Big) du_n \\
&= \int_s^t \frac{(u_n - s)^{n-1}}{(n-1)!} du_n = \frac{(t-s)^n}{n!}.
\end{aligned}
$$

Understanding the iterated integrals - 1d

- Consider again the one-dimensional situation $d = 1$ with X a general continous finite variation path.
- Then we obtain again via induction

$$
\begin{aligned}
\mathbb{X}^{(1)}_{s,t} &= \int_s^t dX_u = X_t - X_s = X_{s,t} \\
\mathbb{X}^{(2)}_{s,t} &= \int_s^t \int_s^{u_2} dX_{u_1} dX_{u_2} = \int_s^t (X_{u_2} - X_s) dX_{u_2} \\
&= \frac{1}{2}(X_t^2 - X_s^2) - X_s(X_t - X_s) = \frac{(X_t - X_s)^2}{2}. \\
&\vdots \\
X^{(n)}_{s,t} &= \int_s^t \int_s^{u_2} \cdots \int_s^{u_n} dX_{u_1} \cdots dX_{u_n} = \int_s^t \Big(\int_s^{u_2} \cdots \int_s^{u_n} dX_{u_1} \cdots dX_{u_{n-1}} \Big) dX_{u_n} \\
&= \int_s^t \frac{(X_{u_n} - X_s)^{n-1}}{(n-1)!} dX_{u_n} = \frac{(X_t - X_s)^n}{(n)!}.
\end{aligned}
$$

Understanding the iterated integrals - 1d

- This is nothing else than the polynomials used in the Taylor expansion.
- By the Taylor formula, a sufficiently regular function $g(X_t)$ can be approximated as a linear function in

$$\mathbb{X}^n_{s,t} = (1, \mathbb{X}^{(1)}_{s,t}, \ldots, \mathbb{X}^{(n)}_{s,t}) = \left(1, X_t - X_s, \ldots, \frac{(X_t - X_s)^n}{(n)!}\right),$$

where the i-th coefficient is given by the i-th derivative $g^{(i)}(X_s)$.

- It is therefore not surprising that the signature serves also more generally as a linear regression basis (on path space).

Understanding the iterated integrals - 2d

- Consider now $d = 2$, for simplicity only up to order $N = 2$.
- Then we get

$$\mathbb{X}^2_{s,t} = \left(1, \begin{pmatrix} \int_s^t dX^1_u \\ \int_s^t dX^2_u \end{pmatrix}, \begin{pmatrix} \int_s^t \int_s^{u_2} dX^1_{u_1} dX^1_{u_2} & \int_s^t \int_s^{u_2} dX^1_{u_1} dX^2_{u_2} \\ \int_s^t \int_s^{u_2} dX^2_{u_1} dX^1_{u_2} & \int_s^t \int_s^{u_2} dX^2_{u_1} dX^2_{u_2} \end{pmatrix}\right).$$

  $\mathbb{X}^2_{s,t}$ thus takes values in $\mathbb{R} \oplus \mathbb{R}^2 \oplus \mathbb{R}^{2\times 2}$.
- For general $d \in \mathbb{N}$ and $N \in \mathbb{N}$ the step-N signature takes values in $\bigoplus_{n=0}^N (\mathbb{R}^d)^{\otimes n}$.

The extended and truncated tensor algebra

- The signature takes values in the extended tensor algebra over $\mathbb{R}^d$ defined by

$$T((\mathbb{R}^d)) := \prod_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n},$$

  with the convention $(\mathbb{R}^d)^{\otimes 0} := \mathbb{R}$.

- Denoting elements of the extended tensor in "blackboard bold face", e.g. $\mathrm{a} = (\mathrm{a}^{(n)})_{n=0}^{\infty} \in T((\mathbb{R}^d))$, $T((\mathbb{R}^d))$ can also be represented as

$$T((\mathbb{R}^d)) = \{\mathrm{a} = (\mathrm{a}^{(0)}, \mathrm{a}^{(1)}, \dots, \mathrm{a}^{(n)}, \dots) | \mathrm{a}^{(n)} \in (\mathbb{R}^d)^{\otimes n} \text{ for all } n \in \mathbb{N}\},$$

The extended and truncated tensor algebra

- The signature takes values in the extended tensor algebra over $\mathbb{R}^d$ defined by

$$T((\mathbb{R}^d)) := \prod_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n},$$

with the convention $(\mathbb{R}^d)^{\otimes 0} := \mathbb{R}$.

- Denoting elements of the extended tensor in "blackboard bold face", e.g. $\mathrm{a} = (\mathrm{a}^{(n)})_{n=0}^{\infty} \in T((\mathbb{R}^d))$, $T((\mathbb{R}^d))$ can also be represented as

$$T((\mathbb{R}^d)) = \{\mathrm{a} = (\mathrm{a}^{(0)}, \mathrm{a}^{(1)}, \dots, \mathrm{a}^{(n)}, \dots) | \mathrm{a}^{(n)} \in (\mathbb{R}^d)^{\otimes n} \text{ for all } n \in \mathbb{N}\},$$

- The step-N signature takes values in the truncated tensor algebra, defined by

$$T^N(\mathbb{R}^d) := \bigoplus_{n=0}^{N} (\mathbb{R}^d)^{\otimes n}.$$

Elements of $T^N(\mathbb{R}^d)$ are denoted in bold face, i.e. $\mathbf{a} = (\mathrm{a}^{(n)})_{n=0}^{N}$.

The extended and truncated tensor algebra

- The signature takes values in the extended tensor algebra over $\mathbb{R}^d$ defined by

$$T((\mathbb{R}^d)) := \prod_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n},$$

with the convention $(\mathbb{R}^d)^{\otimes 0} := \mathbb{R}$.

- Denoting elements of the extended tensor in "blackboard bold face", e.g. $\mathbb{a} = (\mathbb{a}^{(n)})_{n=0}^{\infty} \in T((\mathbb{R}^d))$, $T((\mathbb{R}^d))$ can also be represented as

$$T((\mathbb{R}^d)) = \{\mathbb{a} = (\mathbb{a}^{(0)}, \mathbb{a}^{(1)}, \dots, \mathbb{a}^{(n)}, \dots) | \mathbb{a}^{(n)} \in (\mathbb{R}^d)^{\otimes n} \text{ for all } n \in \mathbb{N}\},$$

- The step-N signature takes values in the truncated tensor algebra, defined by

$$T^N(\mathbb{R}^d) := \bigoplus_{n=0}^{N} (\mathbb{R}^d)^{\otimes n}.$$

Elements of $T^N(\mathbb{R}^d)$ are denoted in bold face, i.e. $\mathbf{a} = (\mathbb{a}^{(n)})_{n=0}^{N}$.

- Elements in $T((\mathbb{R}^d))$ (and $T^N(\mathbb{R}^d)$) can also be written as formal sums, i.e. $\mathbb{a} = \sum_{n=0}^{\infty} \mathbb{a}^{(n)}$ or $\mathbf{a} = \sum_{n=0}^{N} \mathbb{a}^{(n)}$.

Canonical basis elements

- Let $I = (i_1, \dots, i_n)$ be a multi-index with entries in $\{1, \dots, d\}$. The collection of all multi-indices of length n is denoted by $\mathcal{I}_n$. We use the notations,

$$|I| := n, \qquad S(I) := i_1 + i_2 + \cdots + i_n.$$

- Denoting by $\epsilon_1, \dots, \epsilon_d$ the canonical basis of $\mathbb{R}^d$, we use the notations,

$$\epsilon_I := \epsilon_{i_1} \otimes \epsilon_{i_2} \otimes \cdots \otimes \epsilon_{i_n}.$$

- Observe that $(\epsilon_I)_I$ is the canonical orthonormal basis of $(\mathbb{R}^d)^{\otimes n}$.
- Denoting by $\epsilon_\emptyset$ the basis element of $(\mathbb{R}^d)^{\otimes 0}$ we also set $|\emptyset| := 0$.

Some notation and linear functionals

- Let $\pi_n : T((\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes n}$ be the map such that for $\mathrm{a} \in T((\mathbb{R}^d))$, $\pi_n(\mathrm{a}) = \mathrm{a}^{(n)}$, and $\pi_{\leq N} : T((\mathbb{R}^d)) \to T^N(\mathbb{R}^d)$ be such that for $\mathrm{a} \in T((\mathbb{R}^d))$, $\pi_{\leq N}(\mathrm{a}) = \mathbf{a} = (\mathrm{a}^{(n)})_{n=0}^N$.
- Moreover, we introduce the symmetric and antisymmetric parts of $\mathrm{a}^{(2)} \in (\mathbb{R}^d)^{\otimes 2}$:

$$Sym(\mathrm{a}^{(2)}) = \frac{1}{2}(\mathrm{a}^{(2)} + \mathrm{a}^{(2)\,T}), \qquad Anti(\mathrm{a}^{(2)}) = \frac{1}{2}(\mathrm{a}^{(2)} - \mathrm{a}^{(2)\,T}),$$

where, $\mathrm{a}^{(2)\,T}$ denotes the transpose of $\mathrm{a}^{(2)}$.

Some notation and linear functionals

- Let $\pi_n : T((\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes n}$ be the map such that for $\mathrm{a} \in T((\mathbb{R}^d))$, $\pi_n(\mathrm{a}) = \mathrm{a}^{(n)}$, and $\pi_{\leq N} : T((\mathbb{R}^d)) \to T^N(\mathbb{R}^d)$ be such that for $\mathrm{a} \in T((\mathbb{R}^d))$, $\pi_{\leq N}(\mathrm{a}) = \mathbf{a} = (\mathrm{a}^{(n)})_{n=0}^N$.
- Moreover, we introduce the symmetric and antisymmetric parts of $\mathrm{a}^{(2)} \in (\mathbb{R}^d)^{\otimes 2}$:

$$Sym(\mathrm{a}^{(2)}) = \frac{1}{2}(\mathrm{a}^{(2)} + \mathrm{a}^{(2)\,T}), \qquad Anti(\mathrm{a}^{(2)}) = \frac{1}{2}(\mathrm{a}^{(2)} - \mathrm{a}^{(2)\,T}),$$

  where, $\mathrm{a}^{(2)\,T}$ denotes the transpose of $\mathrm{a}^{(2)}$.
- Given $\mathrm{a} \in T((\mathbb{R}^d))$, we write $\mathrm{a}_I := \langle \epsilon_I, \mathrm{a} \rangle$.
- We then define the following set

$$L := \mathrm{span}\{\mathrm{a} \mapsto \mathrm{a}_I \colon |I| \geq 0\},$$

  and call elements of L linear functionals on $T((\mathbb{R}^d))$.

Tensor multiplication

- We equip $T((\mathbb{R}^d))$ and $T^N(\mathbb{R}^d)$ with the standard addition $+$, scalar multiplication and tensor multiplication $\otimes$. In the case of $T^N(\mathbb{R}^d)$ it is truncated at level N.
- For $\mathrm{a}^{(n)} = \sum_{I \in \mathcal{I}_n} a_I \epsilon_I \in (\mathbb{R}^d)^{\otimes n}$ and $\mathrm{b}^{(k)} = \sum_{J \in \mathcal{I}_k} b_J \epsilon_J \in (\mathbb{R}^d)^{\otimes k}$, the tensor multiplication $\mathrm{a}^{(n)} \otimes \mathrm{b}^{(k)} \in (\mathbb{R}^d)^{\otimes (n+k)}$ is defined as follows

$$\mathrm{a}^{(n)} \otimes \mathrm{b}^{(k)} = \sum_{I \in \mathcal{I}_n, J \in \mathcal{I}_k} a_I b_J (\epsilon_I \otimes \epsilon_J).$$

- For $\mathbf{a}, \mathbf{b} \in T^N(\mathbb{R}^d)$, we then have

$$\mathbf{a} \otimes \mathbf{b} = \sum_{n+k \leq N} \mathrm{a}^{(n)} \otimes \mathrm{b}^{(k)},$$

which is equivalent to

$$\pi_m(\mathbf{a} \otimes \mathbf{b}) = \sum_{i=0}^{m} \mathrm{a}^{(m-i)} \otimes \mathrm{b}^{(i)}, \quad \forall m \in \{0, 1, \dots, N\}.$$

Algebra structure of $T^N(\mathbb{R}^d)$

It is thus straightforward to verify the following proposition.

Proposition

The vector space $(T^N(\mathbb{R}^d), +, \cdot)$ *becomes an associative algebra under* $\otimes$ *with neutral element*

$$\mathbf{1} := \epsilon_\emptyset = (1, 0, \ldots, 0) \in T^N(\mathbb{R}^d).$$

Algebra structure of $T^N(\mathbb{R}^d)$

It is thus straightforward to verify the following proposition.

Proposition

The vector space $(T^N(\mathbb{R}^d), +, \cdot)$ *becomes an associative algebra under* $\otimes$ *with neutral element*

$$\mathbf{1} := \epsilon_\emptyset = (1, 0, \dots, 0) \in T^N(\mathbb{R}^d).$$

- Let us now define a norm on $T^N(\mathbb{R}^d)$. For any $\mathbf{a} \in T^N(\mathbb{R}^d)$, we set

$$|\mathbf{a}|_{T^N(\mathbb{R}^d)} := \max_{n=0,\dots,N} |\mathrm{a}^{(n)}|_{(\mathbb{R}^d)^{\otimes n}},$$

  where for $\mathrm{a}^{(n)} = \sum_{I \in \mathcal{I}_n} a_I \epsilon_I$

$$|\mathrm{a}^{(n)}|_{(\mathbb{R}^d)^{\otimes n}} = \sqrt{\sum_{I \in \mathcal{I}_n} |a_I|^2}.$$

- We denote by ρ the relative induced distance, i.e.

$$\rho(\mathbf{a}, \mathbf{b}) := \max_{n=0,\dots,N} |\mathrm{a}^{(n)} - \mathrm{b}^{(n)}|_{(\mathbb{R}^d)^{\otimes n}}, \qquad \mathbf{a}, \mathbf{b} \in T^N(\mathbb{R}^d).$$

The signature ODE

Given a continuous path of bounded variation $X : [0, T] \to \mathbb{R}^d$ and a fixed $s \in [0, T)$, then almost by definition, the path $t \mapsto \mathbb{X}^N_{s,t}$ satisfies an ODE on $T^N(\mathbb{R}^d)$ driven by X.

Proposition

Let $X : [0, T] \to \mathbb{R}^d$ be a continuous path of bounded variation, and let $s \in [0, T)$ be fixed. Then

$$d\mathbb{X}^N_{s,t} = \mathbb{X}^N_{s,t} \otimes dX_t,$$
$$\mathbb{X}^N_{s,s} = \mathbf{1}.$$

Remark: If we define the linear vector fields $U^i : T^N(\mathbb{R}^d) \to T^N(\mathbb{R}^d)$ by $g \mapsto g \otimes \epsilon_i$, then we can rewrite the above ODE as

$$d\mathbb{X}^N_{s,t} = \sum_{i=1}^{d} U^i(\mathbb{X}^N_{s,t}) dX^i_t = \sum_{i=1}^{d} \mathbb{X}^N_{s,t} \otimes \epsilon_i \, dX^i_t.$$

Proof

Consider for $n \geq 1$ the nth-level of the signature given by

$$\begin{aligned}
&\int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dX_{u_1} \otimes \cdots \otimes dX_{u_n} \\
&\quad = \int_s^t \left(\int_s^{u_n} \cdots \int_s^{u_2} dX_{u_1} \otimes \cdots dX_{u_{n-1}} \right) \otimes dX_{u_n} \\
&\quad = \int_s^t \pi_{n-1}(\mathbb{X}^N_{s,u}) \otimes dX_u .
\end{aligned}$$

We thus have

$$\mathbb{X}^N_{s,t} = \mathbf{1} + \int_s^t \mathbb{X}^N_{s,u} \otimes dX_u .$$

Proof

Consider for $n \geq 1$ the nth-level of the signature given by

$$\begin{aligned}\int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dX_{u_1} \otimes \cdots \otimes dX_{u_n} \\ = \int_s^t \left(\int_s^{u_n} \cdots \int_s^{u_2} dX_{u_1} \otimes \cdots dX_{u_{n-1}} \right) \otimes dX_{u_n} \\ = \int_s^t \pi_{n-1}(\mathbb{X}^N_{s,u}) \otimes dX_u.\end{aligned}$$

We thus have

$$\mathbb{X}^N_{s,t} = \mathbf{1} + \int_s^t \mathbb{X}^N_{s,u} \otimes dX_u.$$

Remark: A similar statement holds true on $T((\mathbb{R}^d))$, i.e. the signature satisfies

$$\begin{aligned}d\mathbb{X}_{s,t} &= \mathbb{X}_{s,t} \otimes dX_t, \\ \mathbb{X}_{s,s} &= \mathbf{1}.\end{aligned}$$

Signature under reparametrizations

Proposition

Let $X : [0, T] \to \mathbb{R}^d$ be a continuous path of bounded variation, $\varphi : [0, T] \to [T_1, T_2]$ a non-decreasing surjection, and write $X_t^{\varphi} := X_{\varphi(t)}$ for the reparametrization of X under φ. Then, for all $s, t \in [0, T]$,

$$\mathbb{X}^N_{\varphi(s),\varphi(t)} = \mathbb{X}^{\varphi,N}_{s,t}.$$

Proof: This is a simple consequence of ODE properties.

Chen's theorem

The following theorem shows how the signature of concatenated paths can be computed.

Theorem

Let $X : [0, T] \to \mathbb{R}^d$ be a continuous path of bounded variation and $0 \leq s < t < u \leq T$. Then

$$\mathbb{X}^N_{s,u} = \mathbb{X}^N_{s,t} \otimes \mathbb{X}^N_{t,u}.$$

This is called Chen's relation.

Proof

- We prove this by induction on N.
- For $N = 0$, the equality is just $1 = 1 \otimes 1 = 1$.
- Assume that it holds for N and all $s < t < u$. We now prove that it holds for $N + 1$.

Proof of Chen's theorem (cont.)

- First observe that in $T^{N+1}(\mathbb{R}^d)$,

$$\mathbb{X}_{s,u}^{N+1} = \mathbf{1} + \int_s^u \mathbb{X}_{s,r}^{N+1} \otimes dX_r = \mathbf{1} + \int_s^u \mathbb{X}_{s,r}^{N} \otimes dX_r.$$

  due to the truncation up to level $N+1$.

- Similarly

$$\mathbb{X}_{s,t}^{N+1} \otimes \int_t^u \mathbb{X}_{s,r}^{N} \otimes dX_r = \mathbb{X}_{s,t}^{N} \otimes \int_t^u \mathbb{X}_{s,r}^{N} \otimes dX_r.$$

- Hence, using the induction hyptothesis, splitting $\mathbb{X}_{s,r}^N = \mathbb{X}_{s,t}^N \otimes \mathbb{X}_{t,r}^N$ when $s < t < r < u$, we get

$$\begin{aligned}
\mathbb{X}_{s,u}^{N+1} &= \mathbf{1} + \int_s^u \mathbb{X}_{s,r}^{N} \otimes dX_r = \mathbf{1} + \int_s^t \mathbb{X}_{s,r}^{N} \otimes dX_r + \int_t^u \mathbb{X}_{s,t}^{N} \otimes \mathbb{X}_{t,r}^{N} \otimes dX_r \\
&= \mathbb{X}_{s,t}^{N+1} + \mathbb{X}_{s,t}^{N+1} \otimes \int_t^u \mathbb{X}_{t,r}^{N} \otimes dX_r = \mathbb{X}_{s,t}^{N+1} \otimes (\mathbf{1} + (\mathbb{X}_{t,u}^{N+1} - \mathbf{1})) \\
&= \mathbb{X}_{s,t}^{N+1} \otimes \mathbb{X}_{t,u}^{N+1}.
\end{aligned}$$

Geometric properties

- Consider the example of the step-2 signature with

$$\mathbb{X}^{(2)}_{s,t,(i,j)} = \int_s^t \int_s^{u_2} dX^i_{u_1} dX^j_{u_2}.$$

- Then the product rule $d(X^i X^j) = X^i dX^j + X^j dX^i$ implies that

$$\begin{aligned} Sym(\mathbb{X}^{(2)}_{s,t})^{i,j} = Sym(\mathbb{X}^{(2)}_{s,t})^{j,i} &= \frac{1}{2}\left(\int_s^t (X^i_u - X^i_s) dX^j_u + \int_s^t (X^j_u - X^j_s) dX^i_u \right) \\ &= \frac{1}{2}(X^i_t - X^i_s)(X^j_t - X^j_s) = \frac{1}{2} X^i_{t,s} X^j_{t,s}, \end{aligned}$$

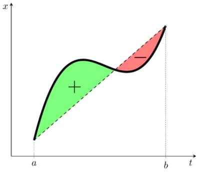
  i.e. for the whole matrix

$$Sym(\mathbb{X}^{(2)}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t}.$$

- This means that the symmetric part of $\mathbb{X}^{(2)}$ is fully determined by $X = \mathbb{X}^{(1)}$.
- To get rid of this redundancy one could only consider $Anti(\mathbb{X}^{(2)})$.

Geometric properties

- Indeed, $Anti(\mathbb{X}^{(2)})$ has an appealing geometric interpretation.
- By definition $Anti(\mathbb{X}_{s,t}^{(2)})^{i,j} = \frac{1}{2}\left(\int_s^t (X_u^i - X_s^i)dX_u^j - \int_s^t (X_u^j - X_s^j)dX_u^i\right)$.
- This is the area (with orientation taken into account) between the curve $\{(X_u^i, X_u^j) : u \in [s,t]\}$ and the chord from (X_s^i, X_s^j) to (X_t^i, X_t^j).



- These properties from first order calculus imply that $T^N(\mathbb{R}^d)$ is actually too big as state space for the signature and that we have to consider a smaller space which has nice geometric properties.

The Lie group $T_1^N(\mathbb{R}^d)$

- Recall that a Lie group is by definition a group which is also a smooth manifold and in which the group operations are smooth maps.

The Lie group $T_1^N(\mathbb{R}^d)$

- Recall that a Lie group is by definition a group which is also a smooth manifold and in which the group operations are smooth maps.
- Set
$$T_c^N(\mathbb{R}^d) := \{\mathbf{a} \in T^N(\mathbb{R}^d) \colon \mathrm{a}_\emptyset = c\}.$$

The Lie group $T_1^N(\mathbb{R}^d)$

- Recall that a Lie group is by definition a group which is also a smooth manifold and in which the group operations are smooth maps.
- Set

$$T_c^N(\mathbb{R}^d) := \{\mathbf{a} \in T^N(\mathbb{R}^d) \colon \mathrm{a}_\emptyset = c\}.$$

The Lie group $T_1^N(\mathbb{R}^d)$

- Recall that a Lie group is by definition a group which is also a smooth manifold and in which the group operations are smooth maps.
- Set

$$T_c^N(\mathbb{R}^d) := \{\mathbf{a} \in T^N(\mathbb{R}^d) \colon \mathrm{a}_\emptyset = c\}.$$

Proposition

$T_1^N(\mathbb{R}^d)$ is a Lie group under the tensor multiplication $\otimes$ (truncated to level N).

The Lie group $T_1^N(\mathbb{R}^d)$

- Proof of the above proposition
  - For any $\mathbf{g}, \mathbf{h} \in T_1^N(\mathbb{R}^d)$, we have $\mathbf{g} \otimes \mathbf{h} \in T_1^N(\mathbb{R}^d)$.
  - As $T^N(\mathbb{R}^d)$ is associative with respect to $\otimes$, this is inheritated by $T_1^N(\mathbb{R}^d)$.
  - The neutral element with respect to $\otimes$ is $\mathbf{1} = \epsilon_\emptyset$.
  - Moreover, for any $\mathbf{a} = (\mathbf{1} + \mathbf{b}) \in T_1^N(\mathbb{R}^d)$, with $\mathbf{b} \in T_0^N(\mathbb{R}^d)$, its inverse is given by

    $$\mathbf{a}^{-1} = \sum_{k=0}^{N} (-1)^k \mathbf{b}^{\otimes k}.$$

    For $N = 2$ we have for example

    $$\mathbf{a}^{-1} = (1, -\mathrm{b}^{(1)}, -\mathrm{b}^{(2)} + \mathrm{b}^{(1)} \otimes \mathrm{b}^{(1)}).$$

  - $T_1^N(\mathbb{R}^d)$ is an affine-linear subspace of $T_N(\mathbb{R}^d)$, hence a smooth manifold. Let us remark that the manifold topology $T_1^N(\mathbb{R}^d)$ is induced by the metric ρ.
  - The group operations $\otimes$ and $^{-1}$ are smooth maps.

The Lie algebra $T_0^N(\mathbb{R}^d)$

- The vector space $T_0^N(\mathbb{R}^d)$ becomes itself an algebra under $\otimes$.
- As in every algebra, the commutator, in our case

$$(\mathbf{g}, \mathbf{h}) \mapsto [\mathbf{g}, \mathbf{h}] := \mathbf{g} \otimes \mathbf{h} - \mathbf{h} \otimes \mathbf{g} \in T_0^N(\mathbb{R}^d)$$

for $\mathbf{g}, \mathbf{h} \in T^N(\mathbb{R}^d)$, defines a bilinear map which
  - is anticommutative, i.e. $[\mathbf{g}, \mathbf{h}] = -[\mathbf{h}, \mathbf{g}]$ and
  - satisfies the Jacobi identity, i.e.

$$[\mathbf{g}, [\mathbf{h}, \mathbf{k}]] + [\mathbf{h}, [\mathbf{k}, \mathbf{g}]] + [\mathbf{k}, [\mathbf{g}, \mathbf{h}]] = 0$$

- Recalling that a vector space V equipped with a bilinear, anticommutative map $[\cdot, \cdot] : V \times V \mapsto V$ which satisfies the Jacobi identity is called a Lie algebra (the map $[\cdot, \cdot]$ is called the Lie bracket), we get ...

Proposition

$(T_0^N(\mathbb{R}^d), +, \cdot, [\cdot, \cdot])$ *is a Lie algebra.*

The exponential and logarithm maps

- To introduce a further Lie group (via the exponential image of a sub Lie-algebra of $T_0^N(\mathbb{R}^d)$) we shall need the notion of the exponential and logarithm maps defined as follows:

$$\exp^{(N)} : T_0^N(\mathbb{R}^d) \to T_1^N(\mathbb{R}^d) \qquad \log^{(N)} : T_1^N(\mathbb{R}^d) \to T_0^N(\mathbb{R}^d)$$

$$\mathbf{b} \mapsto \mathbf{1} + \sum_{k=1}^{N} \frac{\mathbf{b}^{\otimes k}}{k!}, \qquad \mathbf{1} + \mathbf{b} \mapsto \sum_{k=1}^{N} (-1)^{k+1} \frac{\mathbf{b}^{\otimes k}}{k!}.$$

The exponential and logarithm maps

- To introduce a further Lie group (via the exponential image of a sub Lie-algebra of $T_0^N(\mathbb{R}^d)$) we shall need the notion of the exponential and logarithm maps defined as follows:

$$\exp^{(N)}: T_0^N(\mathbb{R}^d) \to T_1^N(\mathbb{R}^d) \qquad \mathbf{b} \mapsto \mathbf{1} + \sum_{k=1}^{N} \frac{\mathbf{b}^{\otimes k}}{k!},$$

$$\log^{(N)}: T_1^N(\mathbb{R}^d) \to T_0^N(\mathbb{R}^d) \qquad \mathbf{1} + \mathbf{b} \mapsto \sum_{k=1}^{N} (-1)^{k+1} \frac{\mathbf{b}^{\otimes k}}{k!}.$$

- For example in the case of $N = 2$ the logarithm is given by

$$\log^{(2)}(\mathbf{1} + \mathbf{b}) = (0, \mathbf{b}^{(1)}, \mathbf{b}^{(2)} - \frac{1}{2}\mathbf{b}^{(1)} \otimes \mathbf{b}^{(1)}).$$

The exponential and logarithm maps

- To introduce a further Lie group (via the exponential image of a sub Lie-algebra of $T_0^N(\mathbb{R}^d)$) we shall need the notion of the exponential and logarithm maps defined as follows:

$$\exp^{(N)}: T_0^N(\mathbb{R}^d) \to T_1^N(\mathbb{R}^d) \qquad \log^{(N)}: T_1^N(\mathbb{R}^d) \to T_0^N(\mathbb{R}^d)$$

$$\mathbf{b} \mapsto \mathbf{1} + \sum_{k=1}^{N} \frac{\mathbf{b}^{\otimes k}}{k!}, \qquad \mathbf{1} + \mathbf{b} \mapsto \sum_{k=1}^{N} (-1)^{k+1} \frac{\mathbf{b}^{\otimes k}}{k!}.$$

- For example in the case of $N = 2$ the logarithm is given by

$$\log^{(2)}(\mathbf{1} + \mathbf{b}) = (0, \mathbf{b}^{(1)}, \mathbf{b}^{(2)} - \frac{1}{2}\mathbf{b}^{(1)} \otimes \mathbf{b}^{(1)}).$$

- A direct calculation shows that $\exp^{(N)}(\log^{(N)}(\mathbf{1} + \mathbf{b})) = \mathbf{b}$, and $\log^{(N)}(\exp^{(N)}(\mathbf{b})) = \mathbf{b}$ for all $\mathbf{b} \in T_0^N(\mathbb{R}^d)$.

The exponential and logarithm maps - Example

- Note that the definitions of $\exp^{(N)}$ and $\log^{(N)}$ are precisely via their classical power series with usual powers replaced by "tensor powers" and the infinite sums replaced by finite ones up to level N .
- Fix some $a \in \mathbb{R}^d$ and consider the path $[0,1] \ni t \mapsto X_t = at$. Then its step-$N$ signature computes as follows

$$\mathbb{X}^N_{0,1} := 1 + \sum_{n=1}^{N} \int_0^1 \int_0^{u_n} \cdots \int_0^{u_2} dX_{u_1} \otimes \cdots \otimes dX_{u_n}$$

$$1 + \sum_{n=1}^{N} a^{\otimes n} \int_0^1 \int_0^{u_n} \cdots \int_0^{u_2} du_1 \cdots du_n$$

$$1 + \sum_{n=1}^{N} \frac{a^{\otimes n}}{n!} = \exp^{(N)}(\mathbf{a}),$$

where $\mathbf{a} = (0, a, 0, \ldots) \in T_0^N$.

The free step-N nilpotent Lie algebra and Lie group

Definition

Define $g^N(\mathbb{R}^d) \subset T_0^N(\mathbb{R}^d)$ as the smallest sub-Lie algebra which contains $\pi_1(T_0^N(\mathbb{R}^d)) = \mathbb{R}^d$. That is,

$$g^N(\mathbb{R}^d) = \mathbb{R}^d \oplus [\mathbb{R}^d, \mathbb{R}^d] \oplus \cdots \oplus [\mathbb{R}^d, [\ldots, [\mathbb{R}^d, \mathbb{R}^d]].$$

We call it the free step-N nilpotent Lie algebra.

By the so-called Campbell-Baker-Hausdorff formula (Theorem 7.26 in Friz & Victoir)

$$\log(\exp(\mathbf{g}) \otimes \exp(\mathbf{h})) \in g^N(\mathbb{R}^d), \quad \mathbf{g}, \mathbf{h} \in g^N(\mathbb{R}^d).$$

It follows that $\exp^{(N)}(g^N(\mathbb{R}^d))$ is a subgroup of $T_1^N(\mathbb{R}^d)$ with respect to $\otimes$.

The free step-N nilpotent Lie algebra and Lie group

Definition

Define $g^N(\mathbb{R}^d) \subset T_0^N(\mathbb{R}^d)$ as the smallest sub-Lie algebra which contains $\pi_1(T_0^N(\mathbb{R}^d)) = \mathbb{R}^d$. That is,

$$g^N(\mathbb{R}^d) = \mathbb{R}^d \oplus [\mathbb{R}^d, \mathbb{R}^d] \oplus \cdots \oplus [\mathbb{R}^d, [\ldots, [\mathbb{R}^d, \mathbb{R}^d]].$$

We call it the free step-N nilpotent Lie algebra.

By the so-called Campbell-Baker-Hausdorff formula (Theorem 7.26 in Friz & Victoir)

$$\log(\exp(\mathbf{g}) \otimes \exp(\mathbf{h})) \in g^N(\mathbb{R}^d), \quad \mathbf{g}, \mathbf{h} \in g^N(\mathbb{R}^d).$$

It follows that $\exp^{(N)}(g^N(\mathbb{R}^d))$ is a subgroup of $T_1^N(\mathbb{R}^d)$ with respect to $\otimes$.

Definition

The image of $g^N(\mathbb{R}^d)$ through the exponential map is a subgroup of $T_1^N(\mathbb{R}^d)$ with respect to $\otimes$. It is called free step-N nilpotent Lie group and is denoted by

$$\exp^{(N)}(g^N(\mathbb{R}^d)).$$

Chow's theorem

- As was seen in the above example, the step-N signature of the path $t \mapsto X_t = at$ for $a \in \mathbb{R}^d$ is precisely $\exp^{(N)}(\mathbf{a}) \in T_1^N(\mathbb{R}^d)$.
- A piecewise linear path, precisely $X : [0, m] \to \mathbb{R}^d$ with $X_i - X_{i-1} = X_{i-1,i} = a_i \in \mathbb{R}^d$, $i = 1, \ldots, m$ for $m \in \mathbb{N}$ and linear between these integer times, is just the concatenation of such paths and by Chen's theorem its step-N signature is of the form

$$\exp^{(N)}(\mathbf{a}_1) \otimes \cdots \otimes \exp^{(N)}(\mathbf{a}_m) \in T_1^N(\mathbb{R}^d)$$

  with $\mathbf{a}_i = (0, a_i, 0, \ldots)$ and $i = 1, \ldots, m$.
- Conversely, any element in $\exp^{(N)}(g^N(\mathbb{R}^d))$ arises as step-N signature of a piecewise linear path of X of the above form (If one prefers, the reparametrization $\widetilde{X}_t = X_{tm}$ defines a piecewise linear path on $[0, 1]$ with identical signature).

Chow's theorem

Theorem

Let $\mathbf{g} \in \exp(g^N(\mathbb{R}^d))$ *. Then, there exist* $a_1, \dots, a_m \in \mathbb{R}^d$ *such that*

$$\mathbf{g} = \exp^{(N)}(\mathbf{a}_1) \otimes \cdots \otimes \exp^{(N)}(\mathbf{a}_m).$$

Equivalently, there exists a piecewise linear path $X : [0,1] \to \mathbb{R}^d$ *with signature* $\mathbf{g}$*, i.e.* $\mathbf{g} = \mathbb{X}^N_{0,1}$.

- This implies that
$$\exp(g^N(\mathbb{R}^d)) = \langle \exp(\mathbb{R}^d) \rangle,$$
where $\langle \exp(\mathbb{R}^d) \rangle = \{\bigotimes_{i=1}^m \exp(\mathbf{a}_i), m \geq 1, \mathbf{a}_i = (0, a_i, 0, \dots), a_i \in \mathbb{R}^d.\}$
- Note that since $g^N(\mathbb{R}^d)$ is closed in $T^N_0(\mathbb{R}^d)$, $\exp(g^N(\mathbb{R}^d))$ is also closed in $T^N_1(\mathbb{R}^d)$.
- Therefore by approximating continuous bounded variation paths by piecewise linear ones it follows that

$$\exp(g^N(\mathbb{R}^d)) = \{\mathbb{X}^N_{0,1} \,|\, \text{signatures of cont. finite variation paths } X\} =: G^N(\mathbb{R}^d).$$

Towards the Carnot-Caratheodory norm

- Chow's theorem tells that for all elements $\mathbf{g} \in G^N(\mathbb{R}^d)$, there exists a continuous path X of finite length such that $\mathbb{X}^N_{0,1} = \mathbf{g}$.
- One may ask for the shortest path (and its length) which has the correct signature.
- For instance, given $a > 0$, we can ask for the shortest path with step-2 signature

$$\exp^{(2)}\left(0 + \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}\right) = (1, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}) \in G^2(\mathbb{R}^2),$$

  or, equivalently, the shortest path in $\mathbb{R}^2$ which ends where it starts and wipes out area a.
- As it is well known the shortest such path is given by a circle (with area a) whose length is given by $2\sqrt{\pi a}$.

The Carnot-Caratheodory norm

Theorem

For every $\mathbf{g} \in G^N(\mathbb{R}^d)$ *, the so-called "Carnot–Caratheodory norm"*

$$\|\mathbf{g}\|_{CC} := \inf\{\int_0^1 |dX_u| : X \in C^{1-var}([0,1], \mathbb{R}^d) \text{ and } \mathbb{X}^N_{0,1} = \mathbf{g}\}$$

is finite and achieved at some minimizing path X^* *, i.e.* $\|\mathbf{g}\|_{CC} = \int_0^1 |dX^*_u|$ *and* $(\mathbb{X}^{*N})_{0,1} = \mathbf{g}$.

Remark

By invariance of length and signatures under reparametrizaion, X^* *need not be defined on* $[0,1]$ *but may be defined for any interval* $[s,t]$ *with non-empty interior.*

Carnot-Caratheodory metric

- The Carnot-Caratheodory norm $\|\cdot\|_{CC}$ induces a metric via

$$d_{CC}(\mathbf{a}, \mathbf{h}) := \|\mathbf{a}^{-1} \otimes \mathbf{h}\|_{CC}, \qquad \mathbf{a}, \mathbf{h} \in G^N(\mathbb{R}^d).$$

- We shall most of the time equip $G^N(\mathbb{R}^d)$ with d_{CC}, making it a metric space.
- The topology on $G^N(\mathbb{R}^d)$ induced by Carnot–Caratheodory distance coincides with the original topology of $G^N(\mathbb{R}^d)$ induced by ρ.

Polynomials of the signature are linear functions

- Consider as example the following identity

$$\langle \epsilon_{(i,i)}, \mathbb{X}_t \rangle = \int_0^t \left(\int_0^s dX_r^i \right) dX_s^i = \int_0^t (X_s^i - X_0^i) dX_s^i = \frac{1}{2}(X_t^i - X_0^i)^2$$
$$= \frac{1}{2} \langle \epsilon_i, \mathbb{X}_t \rangle^2.$$

Polynomials of the signature are linear functions

- Consider as example the following identity

$$\langle \epsilon_{(i,i)}, \mathbb{X}_t \rangle = \int_0^t \left(\int_0^s dX_r^i \right) dX_s^i = \int_0^t (X_s^i - X_0^i) dX_s^i = \frac{1}{2}(X_t^i - X_0^i)^2$$
$$= \frac{1}{2} \langle \epsilon_i, \mathbb{X}_t \rangle^2.$$

- This shows that the quadratic expression on the right hand side has a linear representation.

Polynomials of the signature are linear functions

- Consider as example the following identity

$$\langle \epsilon_{(i,i)}, \mathbb{X}_t \rangle = \int_0^t \left(\int_0^s dX_r^i \right) dX_s^i = \int_0^t (X_s^i - X_0^i) dX_s^i = \frac{1}{2}(X_t^i - X_0^i)^2$$
$$= \frac{1}{2} \langle \epsilon_i, \mathbb{X}_t \rangle^2.$$

- This shows that the quadratic expression on the right hand side has a linear representation.
- This property generalizes to every polynomial function. For the precise statement we first need to introduce a very important operation on the space of multi-indices, namely the shuffle product.

The shuffle product

For a multi-index I denote by $I' = (i_1, \dots, i_{n-1})$.

Definition

For every two multi-indices $I := (i_1, \dots, i_n)$ and $J := (j_1, \dots, j_m)$ the shuffle product is defined recursively as

$$\epsilon_I \sqcup\!\sqcup \epsilon_J := (\epsilon_{I'} \sqcup\!\sqcup \epsilon_J) \otimes \epsilon_{i_n} + (\epsilon_I \sqcup\!\sqcup \epsilon_{J'}) \otimes \epsilon_{j_m},$$

with $\epsilon_I \sqcup\!\sqcup \epsilon_\emptyset := \epsilon_\emptyset \sqcup\!\sqcup e_I = e_I$. It extends to $\mathbf{a}, \mathbf{b} \in T(\mathbb{R}^d)$ as

$$\mathbf{a} \sqcup\!\sqcup \mathbf{b} = \sum_{|I|,|J| \geq 0} a_I b_J (\epsilon_I \sqcup\!\sqcup \epsilon_J).$$

Examples:

- $\epsilon_1 \sqcup\!\sqcup \epsilon_2 = \epsilon_{(\emptyset,1)} \sqcup\!\sqcup \epsilon_{(\emptyset,2)} = \epsilon_{(2,1)} + \epsilon_{(1,2)}$
- $\epsilon_1 \sqcup\!\sqcup \epsilon_{(2,3)} = \epsilon_{(2,3,1)} + \epsilon_{(1,2,3)} + \epsilon_{(2,1,3)}$
- $\epsilon_{(1,2)} \sqcup\!\sqcup \epsilon_{(3,4)} = \epsilon_{(3,4,1,2)} + \epsilon_{(1,3,4,2)} + \epsilon_{(3,1,4,2)} + \epsilon_{(1,3,2,4)} + \epsilon_{(3,1,2,4)} + \epsilon_{(1,2,3,4)}$

The shuffle product property for the signature

Proposition

Let X be continous path of bounded variation and I, J two multi-indices. Then

$$\langle \epsilon_I, \mathbb{X}_{s,t} \rangle \langle \epsilon_J, \mathbb{X}_{s,t} \rangle = \langle \epsilon_I \sqcup\!\sqcup\, \epsilon_J, \mathbb{X}_{s,t} \rangle.$$

Proof: The result follows by induction using integration by parts.

- Fix the multi-index I and let $J = \emptyset$. Then

$$\langle \epsilon_I, \mathbb{X}_{s,t} \rangle \langle \epsilon_\emptyset, \mathbb{X}_{s,t} \rangle = \langle \epsilon_I, \mathbb{X}_{s,t} \rangle = \langle \epsilon_I \sqcup\!\sqcup\, \epsilon_\emptyset, \mathbb{X}_{s,t} \rangle$$

- Now suppose that it holds true for I and J' as well as for I' and J.

The shuffle product property for the signature - Proof continued

- Then by the integration by parts formula

$$
\begin{aligned}
\langle \epsilon_I, \mathbb{X}_{s,t}\rangle \langle \epsilon_J, \mathbb{X}_{s,t}\rangle &= \int_s^t \langle \epsilon_{I'}, \mathbb{X}_{s,u}\rangle dX_u^{i_n} \int_s^t \langle \epsilon_{J'}, \mathbb{X}_{s,u}\rangle dX_u^{j_m} \\
&= \int_s^t \underbrace{\int_s^u \langle \epsilon_{I'}, \mathbb{X}_{s,r}\rangle dX_r^{i_n}}_{\langle \epsilon_I, \mathbb{X}_{s,u}\rangle} \langle \epsilon_{J'}, \mathbb{X}_{s,u}\rangle dX_u^{j_m} \\
&\quad + \int_s^t \underbrace{\int_s^u \langle \epsilon_{J'}, \mathbb{X}_{s,r}\rangle dX_r^{j_m}}_{\langle \epsilon_J, \mathbb{X}_{s,u}\rangle} \langle \epsilon_{I'}, \mathbb{X}_{s,u}\rangle dX_u^{i_n} \\
&= \int_s^t \langle \epsilon_I \sqcup\!\sqcup \epsilon_{J'}, \mathbb{X}_{s,u}\rangle dX_u^{j_m} + \int_s^t \langle \epsilon_J \sqcup\!\sqcup \epsilon_{I'}, \mathbb{X}_{s,u}\rangle dX_u^{i_n} \\
&= \langle \epsilon_I \sqcup\!\sqcup \epsilon_J, \mathbb{X}_{s,t}\rangle.
\end{aligned}
$$

Group-like elements

- We define the set of group-like elements as follows

$$G((\mathbb{R}^d)) := \{\mathrm{a} \in T((\mathbb{R}^d)) \mid \pi_{\leq N}(\mathrm{a}) \in G^N(\mathbb{R}^d) \text{ for all } N\}.$$

- Let $\mathrm{a} \in G((\mathbb{R}^d))$ be a group-like element and $I \in \{1, \ldots, d\}^n$, $J \in \{1, \ldots, d\}^m$ two multi-indices.
- Then, we have as a consequence of Chow's theorem that

$$\langle \epsilon_I, \mathrm{a} \rangle \langle \epsilon_J, \mathrm{a} \rangle = \langle \epsilon_I \sqcup\!\sqcup \epsilon_J, \mathrm{a} \rangle.$$

p-variation norms

- Let (E, d) be a metric space equipped with metric d.
- Let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_k = T\}$ denote again a partition of $[0, T]$. For $p > 0$, we define the p-variation of a path $X \in C([0, T], E)$ by

$$\|X\|_{p-var} := \sup_{\mathcal{D}\subset[0,T]} \left(\sum_{t_i \in \mathcal{D}} d(X_{t_i}, X_{t_{i+1}})^p \right)^{\frac{1}{p}}.$$

- We denote the space of all continuous paths of finite p-variation by $C^p([0, T], E)$.
- As a special case of (E, d) we consider $(G^N(\mathbb{R}^d), d_{CC})$.
- For $\mathbf{X} \in C([0, T], G^N(\mathbb{R}^d))$, we denote the group path increment via $\mathbf{X}_{s,t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$. Consistently with the notation previously used we set

$$\|\mathbf{X}\|_{p-var} := \sup_{\mathcal{D}\subset[0,T]} \left(\sum_{t_i \in \mathcal{D}} d_{CC}(\mathbf{X}_{t_i}, \mathbf{X}_{t_{i+1}})^p \right)^{\frac{1}{p}} = \sup_{\mathcal{D}\subset[0,T]} \left(\sum_{t_i \in \mathcal{D}} \|\mathbf{X}_{t_i, t_{i+1}}\|_{CC}^p \right)^{\frac{1}{p}}.$$

p-variation norms for two-parameter functions

- Additionally, we also consider two-parameter functions $A : \Delta_T \to V$, where $(V, \|\cdot\|)$ is a normed vector space and $\Delta_T := \{(s,t) \in [0,T]^2 \mid s \leq t\}$.
- In this case the *p*-variation is defined as follows

$$\|A\|_{p-var} := \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_i \in \mathcal{D}} \|A_{t_i, t_{i+1}}\|^p \right)^{\frac{1}{p}}.$$

- We stress that if X is a path, then $X_{s,t}$ denotes the increment $X_t - X_s$. Instead, if A is a two-parameter function defined on Δ_T, $A_{s,t}$ denotes the evaluation of A at the pair of times $(s,t) \in \Delta_T$.

Rough paths

Definition

Let $p \in [2,3)$ and $\Delta_T := \{(s,t) \in [0,T]^2 \mid s \leq t\}$. A pair $\mathbf{X} = (X, \mathbb{X}^{(2)})$ is called p-rough path over $\mathbb{R}^d$, in symbols $\mathbf{X} \in \mathcal{C}^p([0,T], \mathbb{R}^d)$, if

$$X : [0,T] \to \mathbb{R}^d, \qquad \mathbb{X}^{(2)} : \Delta_T \to (\mathbb{R}^d)^{\otimes 2}$$

satisfy:

1. The map $[0,T] \ni t \mapsto (X_{0,t}, \mathbb{X}^{(2)}_{0,t}) \in \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2}$ is continuous.
2. Chen's relation holds:
$$\mathbb{X}^{(2)}_{s,u} = \mathbb{X}^{(2)}_{s,t} + \mathbb{X}^{(2)}_{t,u} + X_{s,t} \otimes X_{t,u} \text{ for } 0 \leq s < t < u \leq T.$$
3. $\mathbf{X} = (X, \mathbb{X}^{(2)})$ is of finite p-variation in the rough path sense:
$$|||\mathbf{X}|||_{p-var} := \|X\|_{p-var} + \|\mathbb{X}^{(2)}\|^{1/2}_{p/2-var} < \infty.$$

Some remarks

- $\|X\|_{p-var}$ is the p variation norm for a path with values in $\mathbb{R}^d$ while $\|\mathbb{X}^{(2)}\|_{p/2-var}$ is the p variation distance for a two-parameter function.
- Note that Chen's relation is exactly the same as we got in the signature equation (up to level 2). Indeed, it was given by

$$\mathbb{X}^2_{s,u} = \mathbb{X}^2_{s,t} \otimes \mathbb{X}^2_{t,u} = (1, X_{s,t}, \mathbb{X}^{(2)}_{s,t}) \otimes (1, X_{t,u}, \mathbb{X}^{(2)}_{t,u}),$$

which yields for the second level $\mathbb{X}^{(2)}_{s,u} = \mathbb{X}^{(2)}_{s,t} + \mathbb{X}^{(2)}_{t,u} + X_{s,t} \otimes X_{t,u}$.

- Consider the Lie group valued path $t \mapsto \mathbf{X}_t := (1, X_{0,t}, \mathbb{X}^{(2)}_{0,t}) \in T^2_1(\mathbb{R}^d)$. Define as above the path increments via

$$\mathbf{X}_{s,t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t = (1, -X_{0,s}, -\mathbb{X}^{(2)}_{0,s} + X^{\otimes 2}_{0,s}) \otimes (1, X_{0,t}, \mathbb{X}^{(2)}_{0,t}).$$

and observe that $\mathbf{X}_{s,t} = (1, X_{s,t}, \mathbb{X}^{(2)}_{s,t})$ where $\mathbb{X}^{(2)}_{s,t}$ is given by

$$\mathbb{X}^{(2)}_{s,t} = \mathbb{X}^{(2)}_{0,t} - \mathbb{X}^{(2)}_{0,s} - X_{0,s} \otimes X_{s,t},$$

which is in line with Chen's relation.

Some remarks and weakly geometric rough paths

- Notice that $\mathbf{X}_t = \mathbf{X}_{0,t}$, and for $0 \leq s < t < u \leq 1$ $\mathbf{X}_{s,u} = \mathbf{X}_{s,t} \otimes \mathbf{X}_{t,u}$.
- Hence, this definition of the path increments in $T_1^2(\mathbb{R}^d)$ allows to get intrinsically Chen's relation on the level of the group valued path.

Some remarks and weakly geometric rough paths

- Notice that $\mathbf{X}_t = \mathbf{X}_{0,t}$, and for $0 \leq s < t < u \leq 1$ $\mathbf{X}_{s,u} = \mathbf{X}_{s,t} \otimes \mathbf{X}_{t,u}$.
- Hence, this definition of the path increments in $T_1^2(\mathbb{R}^d)$ allows to get intrinsically Chen's relation on the level of the group valued path.

To mimick the first order calculus, the set of weakly geometric rough paths is introduced as follows:

Definition

Let $p \in [2, 3)$ and $\mathbf{X} \in \mathcal{C}^p([0, T], \mathbb{R}^d)$.
$\mathbf{X}$ is said to be a weakly geometric p-rough path over $\mathbb{R}^d$, in symbols $\mathbf{X} \in \mathcal{C}_g^p([0, T], \mathbb{R}^d)$, if for all $0 \leq s < t \leq 1$

$$Sym(\mathbb{X}_{s,t}^{(2)}) = \frac{1}{2} X_{s,t} \otimes X_{s,t}.$$

Relation to geometric rough path

- There is also the notion of geometric p-rough path, which are precisely limits with respect to the p-variation distance of truncated signatures of order 2 of smooth paths, i.e. a sequence of $(\mathbb{X}^{2,k})_{k\in\mathbb{N}}$ steming from a smooth path X.
- The set of geometric paths is strictly smaller than weakly geometric paths. (The situation is similar to the classical situation of the set of C^p functions being strictly larger than the closure of smooth functions under the p-variation norm).

Weakly geometric rough paths as $G^2(\mathbb{R}^d)$ valued paths

- For a weakly geometric rough path $\mathbf{X}$, it can be deduced that the path $t \mapsto \mathbf{X}_t = (1, X_{0,t}, \mathbb{X}^{(2)}_{0,t}) \in T_1^2(\mathbb{R}^d)$ actually takes values in the $G^2(\mathbb{R}^d) \subset T_1^2(\mathbb{R}^d)$.
- Indeed, recall that $G^2(\mathbb{R}^d) = \exp^{(2)}(g^2(\mathbb{R}^d))$, and $g^2(\mathbb{R}^d) = \mathbb{R}^d \oplus [\mathbb{R}^d, \mathbb{R}^d]$, with $[\mathbb{R}^d, \mathbb{R}^d] := \text{span}\{\epsilon_i \otimes \epsilon_j - \epsilon_j \otimes \epsilon_j,\ 1 \leq i, j \leq d\}$, where $\{\epsilon_i,\ 1 \leq i \leq d\}$ denotes the standard basis of $\mathbb{R}^d$. Thus, $[\mathbb{R}^d, \mathbb{R}^d]$ is nothing but the space of antisymmetric $d \times d$ matrices, and we have that

$$\mathbf{X}_t = (1, X_{0,t}, \frac{1}{2} X_{0,t}^{\otimes 2} + Anti(\mathbb{X}^{(2)}_{0,t})) = \exp^{(2)}(X_{0,t}, Anti(\mathbb{X}^{(2)}_{0,t})) \in G^2(\mathbb{R}^d).$$

- Finally the analytic condition on the p-variation in the definition of a rough path can be equivalently expressed by means of the CC distance d_{CC} on $G^2(\mathbb{R}^d)$.

Weakly geometric rough paths as $G^2(\mathbb{R}^d)$ valued paths

- After the previous remarks, we can adopt a Lie group valued-paths point of view.
- Recall that $T_1^1(\mathbb{R}^d) = G^1(\mathbb{R}^d) = \{1\} \oplus \mathbb{R}^d$.

Definition

Let $p \in [1, 3)$. A continuous path $\mathbf{X} : [0, T] \to G^{[p]}(\mathbb{R}^d) \subset T_1^{[p]}(\mathbb{R}^d)$ is said to be a weakly geometric p-rough path over $\mathbb{R}^d$ if $\|\mathbf{X}\|_{p-var} < \infty$ with $\| \cdot \|_{p-var}$ defined via the CC-distance.

Also for the Lie group valued point of view we denote the set of rough path by $\mathcal{C}^p([0, T), \mathbb{R}^d)$.

Towards signature - Lyons lift

Theorem (Lyons (1998))

Let $p \in [2,3)$ and $\mathbb{N} \ni N > 2$. A weakly geometric p-rough path $\mathbf{X} : [0,T] \to G^2(\mathbb{R}^d)$ admits a unique extension to a path $\mathbb{X}^N : [0,T] \to G^N(\mathbb{R}^d)$, i.e. $\pi_{\leq 2}(\mathbb{X}^N) = \mathbf{X}$, such that

- *$\mathbb{X}^N$ starts from $\mathbf{1} \in G^N(\mathbb{R}^d)$,*
- *it is of finite p-variation, with respect to the Carnot-Caratheodory metric d_{CC} on $G^N(\mathbb{R}^d)$.*

Remark:

- A proof can also be found in Friz & Victoir, Theorem 9.5.

A rough differential equation for the Lyons' lift

Theorem

Let $p \in [2,3)$, $\mathbb{N} \ni N > 2$, and $\mathbf{X} : [0,T] \to G^2(\mathbb{R}^d)$ be a weakly geometric p-rough path. The Lyon's extension $\mathbb{X}^N$ with values in $G^N(\mathbb{R}^d)$ satisfies the following linear rough differential equation (RDE)

$$d\mathbb{X}^N_{s,t} = \mathbb{X}^N_{s,t} \otimes d\mathbf{X}_t,$$
$$\mathbb{X}^N_{s,s} = \mathbf{1} \in G^N(\mathbb{R}^d),$$

which reads in integral form as

$$\mathbb{X}^N_{s,t} = 1 + \int_s^t \mathbb{X}^N_{s,u} \otimes d\mathbf{X}_u,$$

where the integral is understood as rough integral.

We refer to Friz & Victoir for the definition of the rough integral.

Definition of the signature for a p-rough path

As a result the following definition of the signature of a weakly geometric p-rough path follows without ambiguity.

Definition

Let $p \in [2,3)$ and $\mathbf{X} : [0, T] \to G^2(\mathbb{R}^d)$ be a weakly geometric p-rough path. The signature of $\mathbf{X}$, denoted by $\mathbb{X}$, is the unique solution to the RDE in the extended tensor algebra

$$\begin{aligned} d\mathbb{X}_{s,t} &= \mathbb{X}_{s,t} \otimes d\mathbf{X}_t \\ \mathbb{X}_{s,s} &= (1, 0, 0, \dots) \in T((\mathbb{R}^d)). \end{aligned}$$

Semimartingales as rough paths

Continous semimartingales fit well into the theory of rough paths. Indeed, every semimartingale admits a canonical lift which is a.s. a weakly geometric p-rough path for any $p \in (2,3)$.

Proposition

Let $p \in (2,3)$ and X be a continuous $\mathbb{R}^d$-valued semimartingale and $[X,X]^c$ its $(\mathbb{R}^d)^{\otimes 2}$-valued continuous quadratic variation. Then,
$\mathbf{X}(\omega) = (X(\omega), \mathbb{X}^{(2)}(\omega)) \in \mathcal{C}_g^p([0,T], \mathbb{R}^d)$ *a.s., where, for* $0 \leq s \leq t \leq T$,

$$\mathbb{X}^{(2)}_{s,t} := \int_s^t X_{s,r} \otimes dX_r + \frac{1}{2}[X,X]^c_{s,t} = \int_s^t X_{s,r} \otimes \circ dX_r$$

and the first integral is understood in Itô's sense and the second in Stratonovich sense. The lift is called Stratonovich lift.

Signature Stratonovich SDE

Proposition

Let X be a continuous $\mathbb{R}^d$-valued semimartingale and $\mathbf{X}$ its Stratonovich lift. It holds that the above linear RDE for the signature coincides a.s. with the following $T((\mathbb{R}^d))$-valued Stratonovich SDE

$$d\mathbb{X}_{s,t} = \mathbb{X}_{s,t} \otimes \circ dX_t$$
$$\mathbb{X}_{s,s} = (1, 0, 0, \dots) \in T((\mathbb{R}^d)).$$

The explicit solution of this SDE are simply the interated integrals in Stratonovich sense, collected in the following $T((\mathbb{R}^d))$ (or rather $G((\mathbb{R}^d))$) valued object

$$\mathbb{X}_{s,t} = 1 + \int_0^t \mathbb{X}_{s,r} \otimes \circ dX_r,$$

which in coordinate form, for a mulit-index $I = (i_1, \dots, i_n)$, reads as

$$\mathbb{X}^{(n)}_{s,t;I} := \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dX^{i_1}_{u_1} \circ \cdots \circ dX^{i_n}_{u_n} \in \mathbb{R}.$$

Signature of continuous $\mathbb{R}^d$-valued semimartingales

- Hence the signature of an $\mathbb{R}^d$-valued continuous semimartingale X can be defined via

$$\mathbb{X}_{s,t} := \Big(1, \int_s^t \circ dX_s, \int_s^t \int_s^{u_2} \circ dX_{u_1} \otimes \circ dX_{u_2}, \dots, \\ \cdots \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} \circ dX_{u_1} \otimes \cdots \otimes \circ dX_{u_n}, \dots \Big).$$

Towards the universal approximation theorem (UAT)

- Define the following set

$$\widehat{\mathcal{C}}_g^p([0,T],\mathbb{R}^{d+1}) := \{\hat{\mathbf{X}} = (\hat{X}, \hat{\mathbb{X}}^{(2)}) \in \mathcal{C}_g^p([0,T],\mathbb{R}^{d+1}) \mid \text{the first component of } \hat{X} \text{ is } t\}.$$

  We adapt again the Lie group valued-paths point of view to this set. The index of the first component corresponding to t is denoted by -1.

- Consider $C^p([0,T], G^N(\mathbb{R}^d))$ equipped with the p-variation norm defined via the CC-metric defined above, i.e.

$$\|\hat{\mathbf{X}}\|_{p-var} := \sup_{\mathcal{D}\subset[0,T]} \left(\sum_{t_i \in \mathcal{D}} d_{CC}(\hat{\mathbf{X}}_{t_i}, \hat{\mathbf{X}}_{t_{i+1}})^p \right)^{\frac{1}{p}} = \sup_{\mathcal{D}\subset[0,T]} \left(\sum_{t_i \in \mathcal{D}} \|\hat{\mathbf{X}}_{t_i,t_{i+1}}\|_{CC}^p \right)^{\frac{1}{p}}.$$

- From this the following distance is deduced via

$$d_{p-var}(\mathbf{X},\mathbf{Y}) := \sup_{\mathcal{D}\subset[0,T]} \left(\sum_{t_i \in \mathcal{D}} d_{CC}(\mathbf{X}_{t_i,t_{i+1}}, \mathbf{Y}_{t_i,t_{i+1}})^p \right)^{\frac{1}{p}},$$

Universal approximation theorem for continuous functionals of weakly geometric rough paths

Theorem

Let $K \subset \widehat{\mathcal{C}}_g^p([0,T],\mathbb{R}^{d+1})$ be a subset which is compact and let $f : K \to \mathbb{R}$ be continuous, both with respect the above p-variation norm. For each $\hat{\mathbf{X}} \in K$, denote by $\hat{\mathbb{X}}$ its signature. Then, for every $\epsilon > 0$ there exists a linear functional ℓ such that

$$\sup_{\hat{\mathbf{X}}_{[0,T]} \in K} |f(\hat{\mathbf{X}}_{[0,T]}) - \ell(\hat{\mathbb{X}}_{0,T})| \leq \epsilon.$$

Remark: For a version for càdlàg rough paths we refer to C.C, F.Primavera and S.Svaluto-Ferro, 2022.

Proof

- Apply the Stone-Weierstrass theorem to the set A given by

$$A := \text{span}\{\ell : K \to \mathbb{R}\ ;\ \hat{\mathbf{X}} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_{0,T} \rangle : |I| \geq 0\}.$$

- Therefore, we have to prove that A

1. ... is a linear subspace of continuous functions from K to $\mathbb{R}$. This is a consequence of the fact that the Lyons lift

$$(K, d_{p-var}) \to (C^p([0,T], G^N(\mathbb{R}^{d+1})), d_{p-var}),$$
$$\hat{\mathbf{X}} \mapsto \hat{\mathbb{X}}^N$$

is continuous for every $N \geq 3$ (Friz-Victoir,2010, Corollary 9.11). As the evaluation map $\hat{\mathbb{X}}^N \mapsto \hat{\mathbb{X}}^N_{0,T}$ is continuous as well, the claim follows.

Proof

- Apply the Stone-Weierstrass theorem to the set A given by

$$A := \text{span}\{\ell : K \to \mathbb{R}\ ;\ \hat{\mathbf{X}} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_{0,T} \rangle : |I| \geq 0\}.$$

- Therefore, we have to prove that A

1. ... is a linear subspace of continuous functions from K to $\mathbb{R}$. This is a consequence of the fact that the Lyons lift

$$(K, d_{p-var}) \to (C^p([0,T], G^N(\mathbb{R}^{d+1})), d_{p-var}),$$
$$\hat{\mathbf{X}} \mapsto \hat{\mathbb{X}}^N$$

is continuous for every $N \geq 3$ (Friz-Victoir,2010, Corollary 9.11). As the evaluation map $\hat{\mathbb{X}}^N \mapsto \hat{\mathbb{X}}^N_{0,T}$ is continuous as well, the claim follows.

2. ... is a sub-algebra containing a non-zero constant function. This is true by the shuffle-property, as $\hat{\mathbb{X}}^N_{0,T}$ is a group like element.

Proof

- Apply the Stone-Weierstrass theorem to the set A given by

$$A := \text{span}\{\ell : K \to \mathbb{R} \; ; \; \hat{\mathbf{X}} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_{0,T} \rangle : |I| \geq 0\}.$$

- Therefore, we have to prove that A

1. ... is a linear subspace of continuous functions from K to $\mathbb{R}$. This is a consequence of the fact that the Lyons lift

$$(K, d_{p-var}) \to (C^p([0,T], G^N(\mathbb{R}^{d+1})), d_{p-var}),$$
$$\hat{\mathbf{X}} \mapsto \hat{\mathbb{X}}^N$$

is continuous for every $N \geq 3$ (Friz-Victoir,2010, Corollary 9.11). As the evaluation map $\hat{\mathbb{X}}^N \mapsto \hat{\mathbb{X}}^N_{0,T}$ is continuous as well, the claim follows.

2. ... is a sub-algebra containing a non-zero constant function. This is true by the shuffle-property, as $\hat{\mathbb{X}}^N_{0,T}$ is a group like element.

3. ... separates points, which follows from the fact that for a continuous function $f : [0,1] \to \mathbb{R}$ with $f(0) = 0$ and $\int_0^1 f(s)s^n \, ds = 0$ for all $n \in \mathbb{N}$, it holds that $f \equiv 0$.

Proof of point separation

- More precisely, let us consider $\hat{\mathbf{X}}, \hat{\mathbf{Y}} \in K$, with $\hat{\mathbf{X}} \neq \hat{\mathbf{Y}}$.
- Assume by contradiction that their signature is the same.
- Now note that $\int_0^T \hat{X}_s^i \frac{s^n}{n!} ds$ is a linear function of $\hat{\mathbb{X}}_{0,T}$. Indeed it is given by

$$\int_0^T \hat{X}_s^i \frac{s^n}{n!} ds = \langle (\epsilon_i \sqcup\!\sqcup \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \hat{\mathbb{X}}_{0,T} \rangle.$$

- The assumption that the signature of $\hat{\mathbf{X}}, \hat{\mathbf{Y}}$ is the same implies

$$\int_0^T \hat{X}_s^i \frac{s^n}{n!} ds = \int_0^T \hat{Y}_s^i \frac{s^n}{n!} ds,$$

  hence $\hat{X} = \hat{Y}$ by the statement on the previous slide.
- A simlar argument yields $\hat{\mathbb{X}}^{(2)} = \hat{\mathbb{Y}}^{(2)}$ and thus $\hat{\mathbf{X}} = \hat{\mathbf{Y}}$ (for details see C.C, F.Primavera and S.Svaluto-Ferro, 2022). Hence a contradiction.

Remarks

- The above proof shows that the inclusion of time allows to easily show point separation and to avoid so-called tree-like equivalences.
- One essential step is that the Lyons lift is a continuous map from a compact set of rough path with respect to some topology. It works also for the Hölder spaces.
- For applications one crucial point which is often hard to satisfy is the compactness requirement.

$\Rightarrow$ Universal approximation on weighted function spaces where the growth of the functions is controlled by some admissible weight function ψ such that one gets a global approximation result.

Weighted UAT for linear functions of the signature

Without introducing all relevant notation, for the weighted function space

$$\mathcal{B}_{\psi}(\hat{C}^{\alpha}([0,T];G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^{d+1})))$$

where $\hat{C}^{\alpha}([0,T];G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^{d+1}))$ denotes α-Hölder continuous paths with values in $G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^{d+1}))$ one can prove the following global UAT.

Theorem (C.C., P. Schmocker, J. Teichmann ('23))

Let $\psi = \exp(\|\cdot\|^{\gamma}_{CC,\alpha})$ for $\gamma > \lfloor 1/\alpha \rfloor$ be the admissible weight function. The linear span of the set $\left\{\hat{\mathbf{X}} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_T \rangle : I \in \{0,\dots,d\}^n,\ n \in \mathbb{N}\right\}$ is dense in $\mathcal{B}_{\psi}(\hat{C}^{\alpha}([0,T];G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^{d+1})))$, i.e. for every $f \in \mathcal{B}_{\psi}(\hat{C}^{\alpha}([0,T];G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^{d+1})))$ and $\varepsilon > 0$ there exists a linear function ℓ of the signature such that

$$\sup_{\hat{\mathbf{X}}_{[0,T]} \in \hat{C}^{\alpha}} \frac{\left| f(\hat{\mathbf{X}}_{[0,T]}) - \ell(\hat{\mathbb{X}}_{0,T}) \right|}{\psi(\hat{\mathbf{X}}_{[0,T]})} < \varepsilon.$$