Generalized iterated-sums signatures

Nikolas Tapia
joint with J. Diehl and K. Ebrahimi-Fard

FG6

Outline

ISS as feature map

Feature extraction takes a data point $x \in \mathfrak{X}$ and maps it to $\phi(\mathbf{x}) \in \mathfrak{F}$ in *feature space*.

If $\mathfrak{X}$ consists of sequences (time series) in $\mathbb{R}^d$, (Bonnier–Oberhauser–Toth, 2020) propose to define another feature map $\Phi \colon \mathfrak{X} \to \mathrm{T}(\mathfrak{F})$ by

$$\Phi(\mathbf{x}) = \overrightarrow{\prod_{0 \le j < N}} (1 + \phi(\mathbf{x}_j))$$

Depending on the properties of ϕ we get different properties of Φ.

One common choice is the *polynomial augmentation*

$$\phi(\mathbf{x}) = \sum_{n=1}^{\infty} \mathbf{x}^{\otimes n}.$$

The iterated-sums signature (Diehl–Ebrahimi-Fard–T., 2020) is a map from sequence space $\mathfrak{X}$ to a tensor space $\mathrm{T}(V)$, with $V = \mathrm{S}(\mathbb{R}^d)$.

It admits the following factorization (can be taken as def.):

$$\begin{aligned} \mathrm{Sig}(\mathbf{x}) &= \overrightarrow{\prod_{0 \le j < N}} \left(1 + \sum_{n=1}^{\infty} \mathbf{x}_j^{\otimes n}\right) \\ &= \overrightarrow{\prod_{0 \le j < N}} (1 - \mathbf{x}_j)^{-1} \\ &= 1 + \sum_{n=1}^{\infty} \mathrm{Sig}^n(\mathbf{x}) \end{aligned}$$

where

$$\mathrm{Sig}^n(\mathbf{x}) := \sum_{0 \le i_1 < \cdots < i_n < N} \sum_{k_1, \ldots, k_n = 1}^{\infty} \mathbf{x}_{i_1}^{\otimes k_1} \cdots \mathbf{x}_{i_n}^{\otimes k_n}.$$

Quasi-shuffle algebra

Assume we fix a basis $\{e_1, \ldots, e_d\}$ of $\mathbb{R}^d$. Expanding each term we find

$$\mathrm{Sig}^n(\mathbf{x}) = \sum_{\mathcal{I} \in \mathcal{I}_n} M^{\mathcal{I}}(\mathbf{x}) e_{\mathcal{I}}.$$

Here:

- The set $\mathcal{I}$ consists of n-tuples of multi-indices on $\{1, \ldots, d\}$,
- For $\mathcal{I} = (\mathcal{I}_1, \ldots, \mathcal{I}_n) \in \mathcal{I}_n$,

$$e_{\mathcal{I}} = (e_{\mathcal{I}_1^1} \hat{\otimes} \cdots \hat{\otimes} e_{\mathcal{I}_1^{k_1}}) \cdots (e_{\mathcal{I}_n^1} \hat{\otimes} \cdots \hat{\otimes} e_{\mathcal{I}_n^{k_n}}),$$

- For $\mathcal{I} \in \mathcal{I}_n$,

$$M^{\mathcal{I}}(\mathbf{x}) = \sum_{0 \le j_1 < \cdots < j_n < N} \mathbf{x}_{j_1}^{\mathcal{I}_1^1} \cdots \mathbf{x}_{j_1}^{\mathcal{I}_1^{k_1}} \cdots \mathbf{x}_{j_n}^{\mathcal{I}_n^1} \cdots \mathbf{x}_{j_n}^{\mathcal{I}_n^{k_n}}.$$

Shorthand:

$$\begin{aligned} e_{\mathcal{I}} \hat{\otimes} e_{\mathcal{J}} &\sim [\mathcal{I}\mathcal{J}], & \mathcal{I}, \mathcal{J} \in \mathcal{I}_1 \\ e_{\mathcal{I}} e_{\mathcal{J}} &\sim \mathcal{I}\mathcal{J}, & \mathcal{I} \in \mathcal{I}_n, \mathcal{J} \in \mathcal{I}_m \\ e_{\mathcal{I}} &\sim [i_1^1 \cdots i_1^{k_1}] \cdots [i_n \cdots i_n^{k_n}], & \mathcal{I} \in \mathcal{I}_n \\ M^{\mathcal{I}}(\mathbf{x}) &\sim \langle \mathrm{Sig}(\mathbf{x}), e_{\mathcal{I}} \rangle. \end{aligned}$$

Definition (Quasi-shuffle (stuffle, sticky shuffle, ...) product)

For $\mathcal{I} \in \mathcal{I}_n, \mathcal{J} \in \mathcal{I}_m$ and $a, b \in \mathcal{I}_1$,

$$\mathcal{I}a * \mathcal{J}b := (\mathcal{I} * \mathcal{J}b)a + (\mathcal{I}a * \mathcal{J})b + (\mathcal{I} * \mathcal{J})[ab].$$

Example

$$\begin{aligned} i_1[i_2^1 i_2^2] * j_1 = i_1[i_2^1 i_2^2]j_1 + i_1 j_1 [i_2^1 i_2^2] + j_1 i_1 [i_2^1 i_2^2] \\ + [i_1 j_1][i_2^1 i_2^2] + i_1[i_2^1 i_2^2 j_1]. \end{aligned}$$

Some properties

Theorem (Diehl–Ebrahimi-Fard–T., 2020)

The iterated sums-signature satisfies the following properties:

- *Chen's identity:*

$$\mathrm{Sig}(\mathbf{x})_{n,m}\,\mathrm{Sig}(\mathbf{x})_{m,l} = \mathrm{Sig}(\mathbf{x})_{n,l},$$

- *quasi-shuffle identity:*

$$\langle \mathrm{Sig}(\mathbf{x}), I\rangle\langle \mathrm{Sig}(\mathbf{x}), J\rangle = \langle \mathrm{Sig}(\mathbf{x}), I * J\rangle.$$

Theorem (Bonnier–Oberhauser–Toth, 2020)

The iterated-sums signatures is a universal map. That is, under some compactness assumptions, every functional $\mathcal{F} : \mathfrak{X} \to \mathbb{R}$ can be a approximated by a linear function of the iterated-sums signature.

Finer structure of quasi-shuffle

The quasi-shuffle product can be split: $I * J = I \succ J + J \succ I + I \bullet J$ where

$$I \succ Ja = (I * J)a, \quad Ia \bullet Jb = (I * J)[ab].$$

Proposition

The triple $(\mathrm{T}(\mathrm{S}(\mathbb{R}^d)), \succ, \bullet)$ *is a CTD algebra:*

$$I \succ (J \succ K) = (I * J) \succ K,$$
$$(I \succ J) \bullet K = I \succ (J \bullet K),$$
$$(I \bullet J) \bullet K = I \bullet (J \bullet K).$$

In fact, it is the free CTD algebra (Loday, 2007).

Theorem (Diehl–Ebrahimi-Fard–T., 2021)

The iterated-sums signature is the unique CTD morphism such that $i \mapsto (\mathbf{x}^i_k : 0 \leq k < N)$. *In particular*

$$\langle \mathrm{Sig}(\mathbf{x}), I \succ J \rangle = \sum_{j=1}^{N-1} \sum_{i=0}^{j-1} \langle \mathrm{Sig}(\mathbf{x})_{0,i}, I \rangle \langle \mathrm{Sig}(\mathbf{x})_{j,j+1}, J \rangle.$$

Transformations of the first kind

We apply a “formal diffeomorphism” $f \in t\mathbb{R}[[t]]$ on top of the polynomial extension $\phi_P \colon \mathfrak{X} \to \mathrm{S}(\mathbb{R}^d)$.

Let $f(t) = \sum_{n\geq 0} c_n t^n$ with $c_0 = 0$, $c_1 = 1$. Induces a map on the tensor algebra by

$$F(S) = \sum_{n=1}^{\infty} c_n S^n.$$

Let $\Psi_f^* \colon \mathfrak{X} \to \mathrm{T}(\mathrm{S}(\mathbb{R}^d))$ by $\Psi_f^* := F \circ \phi_P$. The resulting signature is

$$\mathrm{Sig}^f(\mathbf{x}) = \overrightarrow{\prod_{0\leq j<N}} \left(1 + F\left(\sum_{n=1}^{\infty} \mathbf{x}_j^{\otimes n}\right)\right)$$

Running example: $f_2(t) = t + \frac{1}{2}t^2$ (in general $f_p(t) = t + \frac{1}{2}t^2 + \cdots + \frac{1}{p!}t^p$ as in Kiraly–Oberhauser, 2019).

$$\mathrm{Sig}^{f_p}(\mathbf{x}) = \overrightarrow{\prod_{0\leq j<N}} \sum_{k=0}^{p} \frac{1}{k!}\left(\sum_{n=1}^{\infty} \mathbf{x}_j^{\otimes n}\right)^k.$$

Transformations of the first kind

A simple case:

$$\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1 \rangle = \sum_{j=0}^{N-1} \mathbf{x}_j^{i_1}$$

$$\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1 i_2 \rangle = \sum_{0 \le j_1 < j_2 < N} \mathbf{x}_{j_1}^{i_1} \mathbf{x}_{j_2}^{i_2} + \frac{1}{2} \sum_{j=0}^{N-1} \mathbf{x}_j^{i_1} \mathbf{x}_j^{i_2}.$$

Therefore,

$$\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1 \rangle \langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_2 \rangle = \langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1 i_2 + i_2 i_1 \rangle.$$

However, it can be checked that this fails for the product

$$\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1 \rangle \langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_2 i_3 \rangle$$

(Hoffman–Ihara, 2017) introduce a map induced by a formal diffeomorphism by performing contractions. For $I = (I_1, \dots, I_n) \in \mathcal{I}_n$ and $\alpha \in \mathcal{C}(n)$ define

$$\alpha[I] = [I_1 \cdots I_{\alpha_1}][I_{\alpha_1+1} \cdots I_{\alpha_1+\alpha_2}] \cdots [I_{\alpha_1+\dots+\alpha_{k-1}+1} \cdots I_n].$$

and let

$$\Psi_f(I) = \sum_{\alpha \in \mathcal{C}(n)} c_{\alpha_1} \cdots c_{\alpha_k} \alpha[I].$$

Example

$(2,1)[i_1 i_2 i_3] = [i_1 i_2] i_3.$

Theorem (Diehl–Ebrahimi-Fard–T., 2021)

The twisted quasi-shuffle

$$I *_f J := \Psi_f^{-1}(\Psi_f(I) * \Psi_f(J))$$

is associative. Moreover, it is also a CTD algebra.

Transformations of the first kind

Theorem (Diehl–Ebrahimi-Fard–T., 2021)

Let f be a formal diffeomorphism. The generalized iterated-sums signature satisfies

$$\langle \mathrm{Sig}^f(\mathbf{x}), I\rangle = \langle \mathrm{Sig}(\mathbf{x}), \Psi_f(I)\rangle,$$

that is,

$$\mathrm{Sig}^f(\mathbf{x}) = \Psi_f^*(\mathrm{Sig}(\mathbf{x})).$$

In particular, $\mathrm{Sig}^f(\mathbf{x})$ *satisfies the twisted quasi-shuffle identity*

$$\langle \mathrm{Sig}^f(\mathbf{x}), I\rangle\langle \mathrm{Sig}^f(\mathbf{x}), J\rangle = \langle \mathrm{Sig}^f(\mathbf{x}), I *_f J\rangle.$$

One can check that, since $f_2^{-1}(t) = \sqrt{2t+1} - 1 = t - t^2/2 + t^3/2 - 5t^4/8 + \cdots$ (and $\Psi_f^{-1} = \Psi_{f^{-1}}$):

$$\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1\rangle\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_2 i_3\rangle = \left\langle \mathrm{Sig}^{f_2}(\mathbf{x}), i_1 i_2 i_3 + i_2 i_1 i_3 + i_2 i_3 i_1 + \frac{1}{2}[i_1 i_2 i_3]\right\rangle$$

The map Ψ_f with $f(t) = e^t - 1$ is known as Hoffman's exponential, and the associated twisted quasi-shuffle is simply the shuffle product.

Transformations of the second kind

Now we only observe a polynomially transformed path: for $P : \mathbb{R}^d \longrightarrow \mathbb{R}^e$ we consider $\mathbf{y} := (P(\mathbf{x}_0), \ldots, P(\mathbf{x}_{N-1}))$.

Example

$P : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $P(\mathbf{x}) = \|\mathbf{x}\|^2$. Then

$$\begin{aligned}\langle \mathrm{Sig}(\mathbf{y}), e_1 e_1 \rangle &= \sum_{0 \le j_1 < j_2 < N} \Big((\mathbf{x}^1_{j_1})^2 + (\mathbf{x}^2_{j_1})^2 \Big) \Big((\mathbf{x}^1_{j_2})^2 + (\mathbf{x}^2_{j_2})^2 \Big) \\ &= \langle \mathrm{Sig}(\mathbf{x}), [11][11] + [11][22] + [22][11] + [22][22] \rangle.\end{aligned}$$

Let $P = (p_1, \ldots, p_e)$, with

$$p_j(\mathbf{x}) = \sum_{\nu \in \mathbb{N}^d} p_{j;\nu} \mathbf{x}^\nu.$$

This induces a map $p_\diamond : \mathbb{R}^e \longrightarrow \mathrm{T}(\mathrm{S}(\mathbb{R}^d))$ by

$$p_\diamond(e_i) = \sum_{\nu \in \mathbb{N}^d} p_{j;\nu}\, e_1^{\hat{\otimes} \nu_1} \,\hat{\otimes} \cdots \hat{\otimes}\, e_d^{\nu_d}.$$

It extends uniquely to a map $\Phi^P : \mathrm{T}(\mathrm{S}(\mathbb{R}^e)) \longrightarrow \mathrm{T}(\mathrm{S}(\mathbb{R}^d))$.

Transformations of the second kind

Theorem (Diehl–Ebrahimi-Fard–T., 2021)

Let $P : \mathbb{R}^d \to \mathbb{R}^e$ be a polynomial transformation, and consider the transformed time series $\mathbf{y} = (P(\mathbf{x}_0), \ldots, P(\mathbf{x}_{N-1}))$. The identity

$$\langle \mathrm{Sig}(\mathbf{y}), I \rangle = \langle \mathrm{Sig}(\mathbf{x}), \Phi^P(I) \rangle$$

holds.

In the previous example: $p_\diamond(e_1) = e_1 \mathbin{\hat{\otimes}} e_1 + e_2 \mathbin{\hat{\otimes}} e_2 \sim [11] + [22]$ so that

$$\begin{aligned} \Phi^P(e_1 e_1) &= p_\diamond(e_1) p_\diamond(e_1) \\ &= ([11] + [22])([11] + [22]) \\ &= [11][11] + [11][22] + [22][11] + [22][22]. \end{aligned}$$

Theorem (Diehl–Ebrahimi-Fard–T., 2021)

The following "Schur-Weyl" duality holds: for any polynomial map P and invertible series f,

$$\Phi^P \circ \Psi_f = \Psi_f \circ \Phi^P.$$

Thanks for your attention!