Multiparameter (iterated) sums

J. **Diehl** (University of Greifswald)
work in progress with L. **Schmitz** (University of Greifswald)

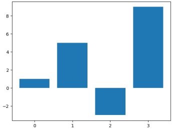
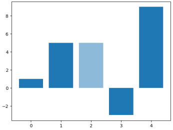
SRA '22

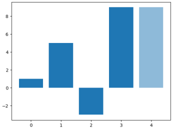
Iterated sums (or integrals) have proven very beneficial in time series analysis.

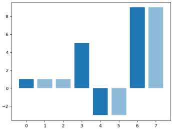
- Graham '13 "Sparse arrays of signatures for . . . ".
- Lyons, Ni, Oberhauser '14 "A feature set for streams . . . "
- various works by L Jin et al '15 on Chinese character recognition.
- Kiraly, Oberhauser '16 "Kernels for sequentially ordered data".
- Lyons, Oberhauser '17 "Sketching the order of events".
- D '13, D,Reizenstein '19 on invariant features.
- D, Ebrahimi-Fard,Tapia '19 "Time warping invariants".
- Kidger, Bonnier, Arribas, Salvi, Lyons '19 "Deep Signature Transforms". - NeurIPS '19
- Kidger, Morrill, Foster, Lyons - Neural Controlled Differential Equations for Irregular Time Series - NeurIPS '20
- Toth, Bonnier, Oberhauser '20 "Seq2Tens".

A key property is their invariance to the parametrization of the data.

For iterated sums this is their time-warping invariance: for example for any iterated sum F,

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

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

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

We are now concerned with two-dimensional data, e.g. images. We would like to have functionals Φ that are warping invariant, for example

$\Phi\Bigg($  $\Bigg) = \Phi\Bigg($  $\Bigg)$

$= \Phi\Bigg($  $\Bigg)$

It is easier to think about invariants to zero-insertion.

$$\Phi(\begin{pmatrix} 0.41 & 0.22 & 0.03 & 0.72 \\ 0.32 & 0.72 & 0.48 & 0.84 \\ 0.60 & 0.64 & 0.25 & 0.93 \end{pmatrix}) = \Phi(\begin{pmatrix} 0.41 & 0.22 & 0 & 0.03 & 0.72 \\ 0.32 & 0.72 & 0 & 0.48 & 0.84 \\ 0.60 & 0.64 & 0 & 0.25 & 0.93 \end{pmatrix})$$

$$= \Phi(\begin{pmatrix} 0.41 & 0.22 & 0 & 0.03 & 0.72 \\ 0.32 & 0.72 & 0 & 0.48 & 0.84 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.60 & 0.64 & 0 & 0.25 & 0.93 \end{pmatrix})$$

(One can afterwards obtain stretching invariants, by plugging in – for example – the discrete second derivative.)

First examples

$$\sum_{i,j} Z_{ij}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j}} Z_{i_1 j} Z_{i_2 j}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_2 j_2}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_1 j_2} Z_{i_2 j_1} Z_{i_2 j_2}$$ Visualization

First examples

$$\sum_{i,j} Z_{ij}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j}} Z_{i_1 j} Z_{i_2 j}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_2 j_2}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_1 j_2} Z_{i_2 j_1} Z_{i_2 j_2}$$ Visualization

First examples

$$\sum_{i,j} Z_{ij}$$ Visualization

$$\sum_{\substack{i_1 < i_2 \\ j}} Z_{i_1 j} Z_{i_2 j}$$ Visualization

$$\sum_{\substack{i_1 < i_2 \\ j_1 < j_2}} Z_{i_1 j_1} Z_{i_2 j_2}$$ Visualization

$$\sum_{\substack{i_1 < i_2 \\ j_1 < j_2}} Z_{i_1 j_1} Z_{i_1 j_2} Z_{i_2 j_1} Z_{i_2 j_2}$$ Visualization

First examples

$$\sum_{i,j} Z_{ij}$$

Visualization

$$\sum_{\substack{i_1<i_2\\ j}} Z_{i_1 j} Z_{i_2 j}$$

Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_2 j_2}$$

Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_1 j_2} Z_{i_2 j_1} Z_{i_2 j_2}$$

Visualization

First examples

$$\sum_{i,j} Z_{ij}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j}} Z_{i_1 j} Z_{i_2 j}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z_{i_1 j_1} Z_{i_2 j_2}$$ Visualization

$$\sum_{\substack{i_1<i_2\\ j_1<j_2}} Z^3_{i_1 j_1} Z^2_{i_1 j_2} Z_{i_2 j_1} Z^4_{i_2 j_2}$$ Visualization

Theorem (D-Schmitz '22)

All polynomial invariants to zero-insertion are given by expressions of the form

$$\left\langle \mathrm{SS}(Z), A \right\rangle := \sum_{\substack{1 \le \iota_1 < \ldots < \iota_m \\ 1 \le \kappa_1 < \ldots < \kappa_n}} \prod_{s=1}^{m} \prod_{t=1}^{n} Z_{\iota_s \kappa_t}^{A_{s\,t}},$$

for *matrix compositions* *A, that is $A \in \mathbb{N}_0^{m \times n}$ containing* <u>*no*</u> *zero rows and* <u>*no*</u> *zero columns.*

Example

$$\left\langle \mathrm{SS}(Z), \begin{bmatrix} 3 \end{bmatrix} \right\rangle = \sum_{i,j} X_{ij}^3 \qquad \left\langle \mathrm{SS}(z), \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\rangle = \sum_{\substack{i_1 < i_2 \\ j}} Z_{i_1 j} Z_{i_2 j}^3$$

$$\left\langle \mathrm{SS}(z), \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \right\rangle = \sum_{\substack{i_1 < i_2 \\ j_1 < j_2}} Z_{i_1 j_1}^3 Z_{i_1 j_2}^2 Z_{i_2 j_1} Z_{i_2 j_2}^4 .$$

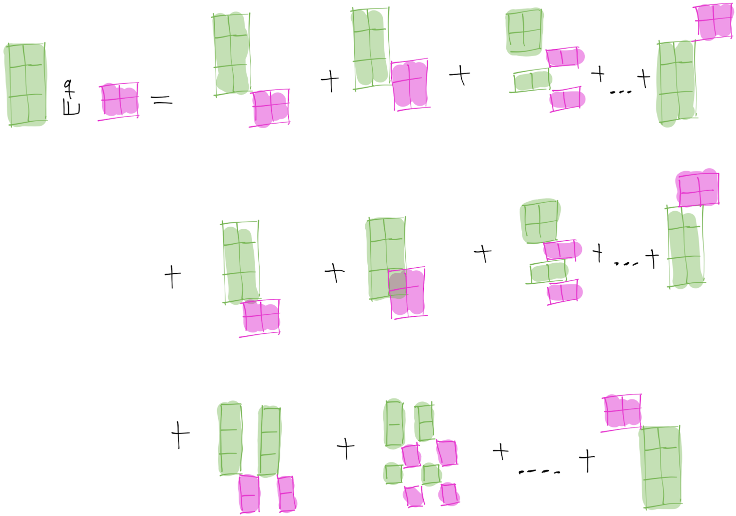
The linear span of matrix compositions becomes an associative, commutative algebra with the quasi-shuffle product $\overset{q}{\sqcup\!\sqcup}$.

Example

$$\begin{bmatrix}1\end{bmatrix} \overset{q}{\sqcup\!\sqcup} \begin{bmatrix}5\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 5\end{bmatrix} + \begin{bmatrix}0 & 1\\ 5 & 0\end{bmatrix} + \begin{bmatrix}5 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}0 & 5\\ 1 & 0\end{bmatrix} + \begin{bmatrix}1\\ 5\end{bmatrix} + \begin{bmatrix}5\\ 1\end{bmatrix} + \begin{bmatrix}1 & 5\end{bmatrix} + \begin{bmatrix}5 & 1\end{bmatrix} + \begin{bmatrix}6\end{bmatrix}$$

$$\begin{bmatrix}1\\ 2\end{bmatrix} \overset{q}{\sqcup\!\sqcup} \begin{bmatrix}3\end{bmatrix} = \begin{bmatrix}1 & 0\\ 2 & 0\\ 0 & 3\end{bmatrix} + \begin{bmatrix}0 & 1\\ 0 & 2\\ 3 & 0\end{bmatrix} + \begin{bmatrix}1\\ 2\\ 3\end{bmatrix} + \begin{bmatrix}1 & 0\\ 0 & 3\\ 2 & 0\end{bmatrix} + \begin{bmatrix}0 & 1\\ 3 & 0\\ 0 & 2\end{bmatrix} + \begin{bmatrix}1\\ 3\\ 2\end{bmatrix} + \begin{bmatrix}0 & 3\\ 1 & 0\\ 2 & 0\end{bmatrix} + \begin{bmatrix}3 & 0\\ 0 & 1\\ 0 & 2\end{bmatrix} + \begin{bmatrix}3\\ 1\\ 2\end{bmatrix} + \begin{bmatrix}1 & 3\\ 2 & 0\end{bmatrix} + \begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix} + \begin{bmatrix}3 & 1\\ 0 & 2\end{bmatrix} + \begin{bmatrix}0 & 1\\ 3 & 2\end{bmatrix} + \begin{bmatrix}4\\ 2\end{bmatrix} + \begin{bmatrix}1\\ 5\end{bmatrix}$$

Quasi-shuffle pictorally



Theorem (Quasi-shuffle identity; D-Schmitz '22)

$$\left\langle \mathsf{SS}(Z), A \right\rangle \cdot_{\mathbb{R}} \left\langle \mathsf{SS}(Z), B \right\rangle = \left\langle \mathsf{SS}(Z), A \stackrel{q}{\sqcup\!\sqcup} B \right\rangle$$

A coproduct would be nice to have for a least two reasons:

1. freeness of the quasi-shuffle algebra,
2. some form of Chen's identity

Currently we have one which gives us point 1) and a weak incarnation of Chen's identity. It is obtained by considering matrix compositions as diagonal 'words' of 'connected' compositions.

Lemma

Every matrix composition can be written as 'diagonal' block matrix of connected matrix compositions which cannot be further decomposed.

Example

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix} = \mathrm{diag}(\begin{pmatrix} 3 \end{pmatrix}, \begin{pmatrix} 5 & 1 \end{pmatrix})$$

Example

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \operatorname{diag}\left(\begin{pmatrix} 3 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 \end{pmatrix}\right)$$

For such a decomposition

$$A = \operatorname{diag}(A_1, \ldots, A_n),$$

we define

$$\Delta A := \sum_{i=0}^{n} \operatorname{diag}(A_1, \ldots, A_i) \otimes \operatorname{diag}(A_{i+1}, \ldots, A_n).$$

Example

$$\Delta \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix} \otimes e + \begin{pmatrix} 3 \end{pmatrix} \otimes \begin{pmatrix} 5 & 1 \end{pmatrix} + e \otimes \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix}$$

Theorem (D-Schmitz '22)

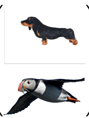
The linear span of matrix compositions with the quasi-shuffle product $\stackrel{q}{\sqcup\!\sqcup}$ *and the 'deconcatenation' coproduct* Δ *is a connected, graded bialgebra (and hence a* *Hopf algebra**).*

SS *is a* *character* *on it. It stores all invariants to zero-insertion. The following form of* *Chen's identity* *holds:*

$\Big\langle \mathrm{SS}\Big($   $\Big), A\Big\rangle = \Big\langle \mathrm{SS}\Big($  $\Big) \otimes \mathrm{SS}\Big($  $\Big), \Delta A\Big\rangle.$

Remark

This is quite a weak form of Chen's identity. It does not help in splitting the calculation of either

$\Big\langle \mathrm{SS}\Big($  $\Big), A\Big\rangle$

or

$\Big\langle \mathrm{SS}\Big($  $\Big), A\Big\rangle$

This is related to the fact that most of the sums we define are not some kind of iterated sum.

Related work

- Chad Giusti, Darrick Lee, Vidit Nanda, and Harald Oberhauser. A topological approach to mapping space signatures, 2022.
- Sheng Zhang, Guang Lin, and Samy Tindel. 2-d signature of images and texture classification, 2022.
- Mohamed R Ibrahim and Terry Lyons. Imagesig: A signature transform for ultra-lightweight image recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3649-3659, 2022.

Summary / Outlook

- Characterization of all zero-insertion invariants, which then lead to (some) warping invariants. ✓
- Stored as a character on a Hopf algebra. ✓
- (Weak) form of Chen's identity. ✓
- By Cartier-Milnor-Moore, the quasi-shuffle algebra is free. Is there a generating set akin to Lynon words ?
- Better Chen's identity / Efficient calculation ?

Question / Comments ?

Summary / Outlook

- Characterization of all zero-insertion invariants, which then lead to (some) warping invariants. ✓
- Stored as a character on a Hopf algebra. ✓
- (Weak) form of Chen's identity. ✓
- By Cartier-Milnor-Moore, the quasi-shuffle algebra is free. Is there a generating set akin to Lynon words ?
- Better Chen's identity / Efficient calculation ?

Question / Comments ?

Complexity

Certain elements, like $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ are cheap ($= \mathcal{O}(T^2)$) to calculate

But most are not. For example,

$$\left\langle \mathsf{SS}(z), \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\rangle$$

seems to be truely of cost $\mathcal{O}(T^3)$.
(Over the Boolean semiring it is $\mathcal{O}(T^\omega)$ ..)