Inverting the signature of a path

Weijun Xu

Joint work with Terry Lyons

November 6, 2014

The signature of a path

Euclidean coordinates: $w = e_{i_1} \cdots e_{i_n}$; define

$$C(w) = \int_{0<u_1<\cdots<u_n<1} d\gamma_{u_1}^{i_1} \cdots d\gamma_{u_n}^{i_n}.$$

The signature of γ is the collection of all $C(w)$'s, denoted by $X(\gamma)$. It captures the ordered evolution along the path through the order of the letters.

Examples:

1. $ax * by \mapsto \exp(ax)\exp(by)$;
2. $by * ax \mapsto \exp(by)\exp(ax)$;
3. $ax + by \mapsto \exp(ax + by)$.

Applications

Crisan-Litterer-Lyons: Nonlinear filtering

Graham: Handwriting recognition, music sound compression.

Ni-Oberhauser: Time series analysis for high frequency data

Papavasiliou-Ladroue: Parameter estimation for rough differential equations

Uniqueness

Hambly-Lyons: paths of finite lengths are uniquely determined by their signatures.

If α, β are two paths of finite lengths, then $X(\alpha) = X(\beta)$ if and only if $\alpha * \beta^{-1}$ is equivalent to a null path.

Question: how to reconstruct the reduced path from its signature?

Boedihardjo-Geng-Lyons-Yang: uniqueness for rough paths.

Main reconstruction theorem

Theorem (Lyons, X.)

By using the signature $X(\gamma)$ only, we can find a piecewise linear path $\tilde{\gamma}$ with k pieces such that

$$\sup_{u\in[0,1]} |\tilde{\gamma}'_u - \gamma'_u| < C\epsilon_k$$

when both are parametrized at unit speed (with respect to ℓ^1 norm), and $\epsilon_k \to 0$ depending on modulus of continuity of γ'.

Commutative in smaller scales; noncommutative in larger scales.

Key: how to rule out noncommutativity in small scales?

Main reconstruction theorem

The piecewise linear path $\tilde{\gamma}$ has the form

$$\tilde{\gamma} = \tilde{\gamma}_1 * \cdots * \tilde{\gamma}_k,$$

where

$$\tilde{\gamma}_j = \frac{\tilde{L}}{k}\left(a_x^{(j)}\rho_j x + a_y^{(j)}(1-\rho_j)y\right).$$

Hope: each $\tilde{\gamma}_j$ approximates $\gamma_{[\frac{j-1}{k},\frac{j}{k}]}$ in the ℓ^1 sense.

$\rho_j, 1-\rho_j \in [0,1]$ represents the unsigned direction;

$a_x^{(j)}, a_y^{(j)} \in \{\pm 1\}$ represents the sign;

$\tilde{L} > 0$ approximates the ℓ^1 length.

Recovering the increment

Symmetrization averages out the order.

Summing over all words of length n with k x's and $n-k$ y's:

$$\mathcal{S}(k, n-k) = \sum_{\mathcal{W}_{k,n-k}} C(w) = \binom{n}{k} (\Delta x)^k (\Delta y)^{n-k}.$$

Maximum: $\frac{k^*}{n-k^*} \approx \frac{|\Delta x|}{|\Delta y|} \quad \Rightarrow \quad$ recovers unsigned direction.

$$\sum_{|k-k^*|<\epsilon} |\mathcal{S}(k, n-k)| \approx \sum_{k} |\mathcal{S}(k, n-k)|$$

Move one level up: comparing $\mathcal{S}(k^* + 1, n-k^*)$ and $\mathcal{S}(k^*, n-k^*)$ gives the sign of the x direction.

Symmetrization

Symmetrizing k blocks with block size $2n$:

$$\underbrace{*****}_{2n} e_{i_1} \underbrace{*****}_{2n} e_{i_2} \cdots\cdots e_{i_{k-1}} \underbrace{*****}_{2n}.$$

Key: pattern in block j are roughly determined by $\gamma_{[\frac{j-1}{k}, \frac{j}{k}]}$.

Steps:

1. Recovering the unsigned directions by checking non-degeneracy.
2. Recovering the signs by moving one level up.
3. Recovering the length by a scaling argument.

Remark: only uses level $2nk + k - 1$ and $2nk + k$.

Summary

Consequences of the reconstruction:

1. Tail signatures already determine $\mathcal{C}^1$ paths.
2. Higher level signatures describe finer structures of the path. Quantitative description?

What have we learned?

1. Symmetrization counts the frequency but neglects the order; so it gives local increments.
2. A certain non-degeneracy criterion is often needed in recovering the directions.