Mathematics for New Technologies in Finance

Solution sheet 3

Through this exercise sheet, we let $E = \mathbb{R}^d$, J an interval on $\mathbb{R}$, and denote $\mathbf{Sig}_J : \mathcal{C}_0^1(J, E) \to \mathbf{T}(E)$ the signature map such that for all $X \in \mathcal{C}_0^1(J, E)$ and we let $\mathbf{Sig}_J^{(M)}$ denote the truncated signature map up to order M: $\mathbf{Sig}_J^{(M)}(X) = (1, \mathbf{s}_1, \cdots, \mathbf{s}_M) \in \mathbf{T}^{(M)}(E)$. Let $X \in \mathcal{C}_0^1([0, s], E)$ and $Y \in \mathcal{C}_0^1([s, t], E)$.

Exercise 3.1 (Signatures)

(a) Let $X_t = t\mathbf{x} \in \mathbb{R}^d$ for all $t \in [0, 1]$. Calculate $\mathbf{Sig}_{[0,1]}(X)$.

(b) Let $X \in \mathcal{C}_0^1([0, T], E)$ and $X_0 = 0$. Prove that

$$\mathbf{Sig}_{[0,1]}(X)_{1,2} + \mathbf{Sig}_{[0,1]}(X)_{2,1} = \mathbf{Sig}_{[0,1]}(X)_1 \cdot \mathbf{Sig}_{[0,1]}(X)_2. \tag{1}$$

Solution 3.1

(a)

$$\mathbf{Sig}_{[0,1]}(X) = (1, \mathbf{x}, \frac{\mathbf{x}^{\otimes 2}}{2!}, \cdots). \tag{2}$$

(b) By integration by part, we directly show the equality

$$\int_0^1 u_t^{(1)} du_t^{(2)} + \int_0^1 u_t^{(2)} du_t^{(1)} = \int_0^1 d(u^{(1)} \cdot u^{(2)})_t = u_1^{(1)} \cdot u_1^{(2)} \tag{3}$$

Exercise 3.2 (Calculate Signatures)

(a) Let $X \in \mathcal{C}_0^1([0, 1], \mathbb{R})$ s.t. $X_t = \sin(t)$ for all $t \in [0, 1]$. Calculate $\mathbf{Sig}_{[0,1]}^{(2)}(X)$ i.e. the signatures of X up to order 2.

(b) Let $X \in \mathcal{C}_0^1([0, 1], \mathbb{R}^2)$ s.t. $X_t = (t, \sin(t))$ for all $t \in [0, 1]$. Calculate $\mathbf{Sig}_{[0,1]}^{(2)}(X)$ i.e. the signatures of X up to order 2.

(c) Let $X \in \mathcal{C}_0^1([0, 1], \mathbb{R})$ and $n \in \mathbb{N}$. Calculate $\int_0^1 t^n dX_t$ when

  (i) $X_t = t$

  (ii) $X_t = \sin(t)$

(d) Prove that

$$\mathcal{F} = \left\{ \mathcal{C}_0^1([0, 1], \mathbb{R}) \ni X \mapsto \sum_{i=1}^n \lambda_i \int t^i dX_t \in \mathbb{R} : \forall \lambda_i \in \mathbb{R}, n \in \mathbb{N} \right\}$$

is a point-separating vector space. $\mathcal{C}_0^1([0, 1], \mathbb{R})$ is the space of all function f on $[0, 1]$ with $f(0) = 0$ and f has continuous derivative.

Solution 3.2

(a)

$$\Big(1, \sin(1), \int_0^1 \sin(t)\cos(t)dt\Big) \tag{4}$$

(b)

$$\Big(1, 1, \sin(1), \frac{1}{2}, \int_0^1 \sin(t)dt, \int_0^1 t\cos(t)dt, \int_0^1 \sin(t)\cos(t)dt\Big) \tag{5}$$

(c) (i)

$$\frac{t^{n+1}}{n+1}\Big|_0^1 \tag{6}$$

(ii)

$$\begin{aligned}\int_0^1 t^n d\sin(t) &= \sin(t)t^n\Big|_0^1 + \int_0^1 nt^{n-1}d\cos(t)\\ &= \sin(t)t^n\Big|_0^1 + \int_0^1 nt^{n-1}d\cos(t)\\ &= \sin(t)t^n\Big|_0^1 + n\cos(t)t^{n-1}\Big|_0^1 - \int_0^1 n(n-1)t^{n-2}d\sin(t)\\ &= \dots\end{aligned} \tag{7}$$

(d) Vector space holds directly from the definition. So we remain to show point-separating. Let us consider $Z \in \mathcal{C}_0^1([0,1], \mathbb{R})$ s.t.

$$\int \sum_{i=1}^n \lambda_i t^i dZ_t = 0, \quad \forall \lambda_i \in \mathbb{R}, n \in \mathbb{N}.$$

An elementary approach is using universal approximation of polynomials. Since Z' is continuous on $[0,1]$, it can be universally approximated by polynomials, and therefore we have

$$\int_0^1 (Z_t')^2 dt = \lim_{n\to\infty} \int \sum_{i=1}^n \lambda_i t^i dZ_t = 0. \tag{8}$$

This implies that $Z = 0$ because it starts from 0, which completes the proof.

Remark: It worth noticing that this essentially relies on that Z' is continuous. But we can actually make the proof more general by considering function X which are only L-Lipschitz and starting from 0, and then a more general proof can be done by fourier analysis. Since $\sin(m\pi t)$ and $\cos(m\pi t)$ for all $m \in \mathbb{N}$ are uniformly approximated by polynomial on $[0,1]$. We have for all $m \in \mathbb{N}$

$$\int \sin(mt)dZ_t = \int \cos(mt)dZ_t = 0 \tag{9}$$

Then we define a sign measure $\mu(dt) = Z_t'dt$ (Because by Rademacher's Lipschitz function is almost everywhere differentiable and here we even know that $|Z_t'| \le L$ almost surely), then for all $m \in \mathbb{N}$

$$\int \sin(mt)d\mu = \int \cos(mt)d\mu = 0. \tag{10}$$

Then by fourier analysis we know $\mu = 0$ so Z is constant, which is actually 0 because $Z(0) = 0$. This proof uses the same idea used in the proof of universal approximation theory of neural network by G. Cybenko.

Exercise 3.3 (Controlled ODEs) Consider the controlled ODE: $X_0 = x \in \mathbb{R}$

$$dX_t^\theta = V^\theta(t, X_t^\theta)dt, \quad t \in [0, T]. \tag{11}$$

(a) Let

$$a_t = \frac{\partial X_T^\theta}{\partial X_t^\theta}. \tag{12}$$

Prove that

$$\frac{d}{dt}a_t = -\frac{\partial V^\theta}{\partial x}(t, X_t^\theta) \cdot a_t, \quad a_T = 1, \tag{13}$$

and relate a_t with $J_{t,T}$ in the lecture notebook.

(b) Prove that

$$\frac{d}{dt}(\frac{\partial X_t^\theta}{\partial \theta}a_t) = a_t\frac{\partial V^\theta}{\partial \theta}(t, X_t^\theta), \tag{14}$$

and

$$\frac{\partial X_T^\theta}{\partial \theta} = -\int_T^0 \frac{\partial X_T^\theta}{\partial X_t^\theta} \cdot \frac{\partial V^\theta}{\partial \theta}(t, X_t^\theta)dt. \tag{15}$$

(c) Is every feedforward neural network a discretization of controlled ODE?

Solution 3.3

(a) We know

$$\begin{aligned} a_t = \frac{\partial X_T^\theta}{\partial X_t^\theta} &= \frac{\partial X_T^\theta}{\partial X_{t+\Delta t}^\theta} \cdot \frac{\partial X_{t+\Delta t}^\theta}{\partial X_t^\theta} \\ &= a_{t+\Delta t} \cdot \frac{\partial X_{t+\Delta t}^\theta}{\partial X_t^\theta}. \end{aligned} \tag{16}$$

Also we know

$$X_{t+\Delta t}^\theta = X_t^\theta + \int_t^{t+\Delta t} V^\theta(X_s^\theta, s)ds \tag{17}$$

Taking partial derivative on both side we have

$$\frac{\partial X_{t+\Delta t}^\theta}{\partial X_t^\theta} = 1 + \int_t^{t+\Delta t} \partial_x V^\theta(X_s^\theta, s)ds \tag{18}$$

Plug this into (16) we have

$$\frac{a_t - a_{t+\Delta t}}{a_{t+\Delta t}} = \int_t^{t+\Delta t} \partial_x V^\theta(X_s^\theta, s)ds. \tag{19}$$

Let $\Delta t \to 0$ we obtain

$$\frac{d}{dt}a_t = -\frac{\partial V^\theta}{\partial x}(t, X_t^\theta) \cdot a_t \tag{20}$$

(b)

$$\begin{aligned} \frac{d}{dt}(\frac{\partial X_t^\theta}{\partial \theta}a_t) &= \frac{d}{dt}(\frac{\partial X_t^\theta}{\partial \theta}) \cdot a_t + \frac{da_t}{dt} \cdot (\frac{\partial X_t^\theta}{\partial \theta}) \\ &= \frac{\partial}{\partial \theta}V^\theta(X_t^\theta, t) \cdot a_t - \frac{\partial V^\theta}{\partial x}(t, X_t^\theta) \cdot a_t \cdot (\frac{\partial X_t^\theta}{\partial \theta}) \\ &= a_t\frac{\partial V^\theta}{\partial \theta}(t, X_t^\theta). \end{aligned} \tag{21}$$

The last equation is because:

$$\frac{\partial}{\partial\theta}V^\theta(X_t^\theta,t) = \frac{\partial V^\theta}{\partial x}(t,X_t^\theta)\cdot(\frac{\partial X_t^\theta}{\partial\theta}) + \frac{\partial V^\theta}{\partial\theta}(t,X_t^\theta). \tag{22}$$

(c) Yes

Exercise 3.4 (Linear controlled ODE) Let $E = \mathbb{R}^d, W = \mathbb{R}^n$. Let $X \in \mathcal{C}_0^1([0,T],E)$ and let $B\colon E \to \mathbf{L}(W)$ be a bounded linear map. Consider

$$dY_t = B(dX_t)(Y_t) \tag{23}$$

If we denote $B^k = B(e_k)$, $k = 1,\cdots,d$ then

$$dY_t = \sum_{k=1}^d B^k(Y_t)dX_t^k. \tag{24}$$

Prove that

$$Y_t = \Big(\sum_{k=0}^\infty B^{\otimes k}\Big)\Big(\mathbf{Sig}_{[0,t]}(X)\Big)Y_0. \tag{25}$$

This implies that the solution of controlled SDE could be written as a linear function on signature stream of driving path. This implies that signature stream is a promising feature for controlled ODE.

Solution 3.4 It follows from Picard's iteration that

$$\begin{aligned} Y_t^n &= \left(I + \sum_{k=1}^n B^{\otimes k}\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}\otimes\cdots\otimes dX_{t_k}\right)Y_0 \\ &= \left(I + \sum_{k=1}^n \sum_{i_1,\cdots,i_k=1}^d B^{i_k}\cdots B^{i_1}\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}^{i_1}\cdots dX_{t_k}^{i_k}\right)Y_0. \end{aligned} \tag{26}$$

Let the variation of $X \in \mathcal{C}_0^1([0,T],E)$ denoted by $\|X\|_{[0,T]}$, then

$$\Big\|\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}\otimes\cdots\otimes dX_{t_k}\Big\|_{E^{\otimes k}} \le \frac{\|X\|_{[0,T]}^k}{k!}. \tag{27}$$

Therefore, Y_t^n converges to Y_t as $n\to\infty$ i.e.

$$\|Y_t - Y_t^n\|_W \le \sum_{k>n}\frac{\|B\|_{\mathcal{L}(E,\mathcal{L}(W))}^k\|X\|_{[0,T]}^k}{k!} \le \frac{\|B\|_{\mathcal{L}(E,\mathcal{L}(W))}^{n+1}\|X\|_{[0,T]}^{n+1}}{n!} \to 0, \quad \text{as } n\to\infty \tag{28}$$

and

$$Y_t = \left(I + \sum_{k=1}^\infty B^{\otimes k}\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}\otimes\cdots\otimes dX_{t_k}\right)Y_0. \tag{29}$$

In the language of signature, we have that

$$Y_t = \Big(\sum_{k=0}^\infty B^{\otimes k}\Big)\Big(\mathbf{Sig}_{[0,t]}(X)\Big)Y_0. \tag{30}$$

This implies that the solution of controlled SDE could be written as a linear function on signature stream of driving path. This implies that signature stream is a promising feature for controlled ODE.

References

[1] Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788*, 2016.

[2] Terry J Lyons, Michael Caruana, and Thierry Lévy. *Differential equations driven by rough paths*. Springer, 2007.