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 $\mathbf{Sig}_J \colon \mathcal{C}_0^1(J, E) \to \mathbf{T}(E)$ the signature map 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: $\mathbf{Sig}_J^{(M)}(X) = (1, \mathbf{s}_1, \cdots, \mathbf{s}_M) \in \mathbf{T}^{(M)}(E)$.

Exercise 3.1 (Controlled ODEs) Consider the controlled ODE: $X_0 = x \in \mathbb{R}$

$$dX_t^\theta = V^\theta(t, X_t^\theta)dt, \quad t \in [0, T]. \tag{1}$$

(a) Let

$$a_t = \frac{\partial X_T^\theta}{\partial X_t^\theta}. \tag{2}$$

Prove that

$$\frac{d}{dt} a_t = -\frac{\partial V^\theta}{\partial x}(t, X_t^\theta) \cdot a_t, \quad a_T = 1, \tag{3}$$

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

(b) Prove that

$$\frac{d}{dt}\left(\frac{\partial X_t^\theta}{\partial \theta} a_t\right) = a_t \frac{\partial V^\theta}{\partial \theta}(t, X_t^\theta), \tag{4}$$

and

$$\frac{\partial X_T^\theta}{\partial \theta} = -\int_T^0 \frac{\partial X_T^\theta}{\partial X_t^\theta} \cdot \frac{\partial V^\theta}{\partial \theta}(t, X_t^\theta)dt. \tag{5}$$

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

Solution 3.1

(a) We know

$$\begin{aligned} a_t = \frac{\partial X_T^\theta}{\partial X_t^\theta} &= \frac{\partial X_T^\theta}{\partial X_{t+\Delta t}^\theta} \cdot \frac{\partial X_{t+\Delta t}^\theta}{\partial X_t^\theta} \\ &= a_{t+\Delta t} \cdot \frac{\partial X_{t+\Delta t}^\theta}{\partial X_t^\theta}. \end{aligned} \tag{6}$$

Also we know

$$X_{t+\Delta t}^\theta = X_t^\theta + \int_t^{t+\Delta t} V^\theta(X_s^\theta, s)ds \tag{7}$$

Taking partial derivative on both side we have

$$\frac{\partial X_{t+\Delta t}^\theta}{\partial X_t^\theta} = 1 + \int_t^{t+\Delta t} \partial_x V^\theta(X_s^\theta, s)ds \tag{8}$$

Plug this into (6) we have

$$\frac{a_t - a_{t+\Delta t}}{a_{t+\Delta t}} = \int_t^{t+\Delta t} \partial_x V^\theta(X_s^\theta, s)ds. \tag{9}$$

Let $\Delta t \to 0$ we obtain

$$\frac{d}{dt} a_t = -\frac{\partial V^\theta}{\partial x}(t, X_t^\theta) \cdot a_t \tag{10}$$

(b)

$$\begin{aligned}\frac{d}{dt}(\frac{\partial X_t^\theta}{\partial \theta}a_t) &= \frac{d}{dt}(\frac{\partial X_t^\theta}{\partial \theta})\cdot a_t + \frac{da_t}{dt}\cdot(\frac{\partial X_t^\theta}{\partial \theta}) \\ &= \frac{\partial}{\partial\theta}V^\theta(X_t^\theta,t)\cdot a_t - \frac{\partial V^\theta}{\partial x}(t,X_t^\theta)\cdot a_t\cdot(\frac{\partial X_t^\theta}{\partial\theta}) \\ &= a_t\frac{\partial V^\theta}{\partial\theta}(t,X_t^\theta).\end{aligned} \tag{11}$$

(c) Yes

Exercise 3.2 (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

$$dY_t = B(dX_t)(Y_t) \tag{12}$$

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

$$dY_t = \sum_{k=1}^d B^k(Y_t)dX_t^k. \tag{13}$$

Prove that

$$Y_t = \Big(\sum_{k=0}^\infty B^{\otimes k}\Big)\Big(\mathbf{Sig}_{[0,t]}(X)\Big)Y_0. \tag{14}$$

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 3.2 It follows from Picard's iteration that

$$\begin{aligned}Y_t^n &= \left(I + \sum_{k=1}^n B^{\otimes k}\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}\otimes\cdots\otimes dX_{t_k}\right)Y_0 \\ &= \left(I + \sum_{k=1}^n \sum_{i_1,\cdots,i_k=1}^d B^{i_k}\cdots B^{i_1}\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}^{i_1}\cdots dX_{t_k}^{i_k}\right)Y_0.\end{aligned} \tag{15}$$

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

$$\Big\|\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}\otimes\cdots\otimes dX_{t_k}\Big\|_{E^{\otimes k}} \le \frac{\|X\|_{[0,T]}^k}{k!}. \tag{16}$$

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

$$\|Y_t - Y_t^n\|_W \le \sum_{k>n}\frac{\|B\|_{\mathcal{L}(E,\mathcal{L}(W))}^k\|X\|_{[0,T]}^k}{k!} \le \frac{\|B\|_{\mathcal{L}(E,\mathcal{L}(W))}^{n+1}\|X\|_{[0,T]}^{n+1}}{n!} \to 0, \quad \text{as } n\to\infty \tag{17}$$

and

$$Y_t = \left(I + \sum_{k=1}^\infty B^{\otimes k}\int_{t_1<\cdots<t_k\in[0,t]} dX_{t_1}\otimes\cdots\otimes dX_{t_k}\right)Y_0. \tag{18}$$

In the language of signature, we have that

$$Y_t = \Big(\sum_{k=0}^\infty B^{\otimes k}\Big)\Big(\mathbf{Sig}_{[0,t]}(X)\Big)Y_0. \tag{19}$$

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] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

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

[3] Yann LeCun, D Touresky, G Hinton, and T Sejnowski. A theoretical framework for back-propagation. In *Proceedings of the 1988 connectionist models summer school*, volume 1, pages 21–28, 1988.

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