Rough Paths:

- Motivación original ← estudiar "ecuaciones diferenciales" que incluyen perturbaciones aleatorias.

- Motivacion reciente ← Realizar aprendizaje de máquina interpretando datos como trayectorias de funciones continuas.

Motivación clásica:

$Y_t = f(t)$ ← Función del tiempo.

$$\frac{dY_t}{dt} = V(Y_t) \quad \text{con} \quad Y_0 = y_0 \qquad V: \mathbb{R} \to \mathbb{R}.$$

$$dY_t = V(Y_t)\,dt \quad \text{con} \quad Y_0 = y_0$$

$$\Updownarrow$$

$$Y_t - Y_s = \int_s^t V(Y_r)\,dr$$

Siguiente paso de generalidad:

$$dY_t = V(Y_t)\, dX_t$$

$X_t \leftarrow$ otra función del tiempo que satisface ciertas condiciones.

$$\Updownarrow$$

$$Y_t - Y_s = \int_s^t V(Y_r)\, dX_r$$

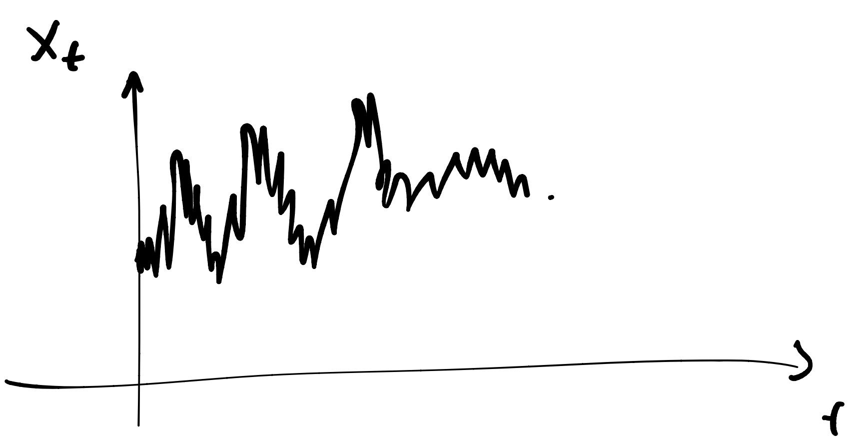
$$\underset{\text{Output}}{dY_t} = V(Y_t)\, \underset{\text{input.}}{dX_t}$$

¿Aplicación? X_t = precio en pesos del dolar.

Y_t = Cartera de inversión de divisas.

Lado negativo:

Si vemos el mercado en tiempo real:



$\leftarrow$ Podría ser un problema.

Análisis detallado del problema:

$$dY_r = V(Y_r)\,dX_r \iff Y_{r+\varepsilon} - Y_r \approx V(Y_r)\,(X_{r+\varepsilon} - X_r).$$

Samplear X en $[0,1]$ en tiempos

$t_1 = \frac{1}{n}$, $t_2 = \frac{2}{n}, \dots,$ $t_n = 1 = \frac{n}{n}$.

Si $X_i = X_{t_i}$ $Y_i = Y_{t_i}$

$$Y_{i+1} = Y_i + V(Y_i)\,\Delta_i X$$

$\Delta X_t \approx \Delta t$ + error

$$Y_{i+1} = Y_i + V(Y_i)\left(\frac{1}{n} + \frac{1}{\sqrt{n}}\,\xi_i\right).$$

modelo

$\{\xi_i\} \leftarrow$ v.a.i.id. $\xi_i \sim \mathcal{N}(0,1)$.

Cuando tomamos límite obtenemos

$$dY_t = V(Y_t)\,(dt + dB_t).$$

$$Y_t - Y_s = \int_s^t V(Y_r)\,dr + \underbrace{\int_s^t V(Y_r)\,dB_r}_{\text{es "sutil" de definir.}}$$

Esto motiva a formular ecuaciones del tipo

$$dY_t = V(Y_t)\, dX_t .$$

$X_t \leftarrow$ Potencialmente tiene trayectorias muy erráticas.

¿Cómo depende el output en función del input?

$\updownarrow$

¿continuamente?

$$dY_t = V(Y_t)\, dX_t .$$

Proposición (Lyons et. al).

No existe un espacio de Banach $\mathcal{B}$ t.q. las trayectorias de un mov. Browniano pertenezcan a $\mathcal{B}$ casi seguramente y t.g. el mapeo

$$(f, g) \longmapsto \int_0^{(\cdot)} f(t)\dot{g}(t)\, dt,$$

definido primero sobre f, g diferenciables $f, g : \mathbb{R} \to \mathbb{R}$ se extiende a $\mathcal{B} \times \mathcal{B}$ de manera continua.

Como esquivar esquivar esta restricción:

$$Y_t - Y_s = \int_s^t V(Y_r)\, dX_r \quad *$$

Framework. general:

$$Y_t = (Y_t^1, \dots, Y_t^e) \qquad e \in \mathbb{N}.$$

$$X_s = (Y_s^1, \dots, Y_s^d). \qquad d \in \mathbb{N}.$$

$$V \in \mathbb{R}^{e \times d} \qquad V = (V_{i,k}\ ; \quad 1 \leq i \leq e, \quad 1 \leq k \leq d).$$

$$Y_t^i - Y_s^i = \sum_{k=1}^{d} \int_s^t V_{i,k}(Y_r)\, dX_r^k \quad \leftarrow \text{nuestra ecuación}$$

se suponer que tengo acceso a Y_s, X_s.

$$Y_t^i - Y_s^i = \sum_{k=1}^{d} \int_s^t \left(V_{i,k}(Y_s) + \left(V_{i,k}(Y_r) - V_{i,k}(Y_s) \right) \right) dX_r^k$$

$$= \sum_{k=1}^{d} V_{i,k}(Y_s) \underbrace{\int_s^t dX_r^k}_{\text{Importante}}$$

$$+ \sum_{k=1}^{d} \int_s^t \sum_{i_2=1}^{e} \int_s^r \frac{\partial V_{i,k}}{\partial y_{i_2}}(Y_{r_2})\, dY_{r_2}^{i_2}\, dX_r^k$$

$$\int_s^t \int_s^r \frac{\partial V_{i,r}}{\partial y_{i_2}}(Y_{r_2})\, dY_{r_2}^{i_2}\, dX_1^K$$

$$= \int_s^t \int_s^r \frac{\partial V_{i,r}}{\partial y_{i_2}}(Y_{r_2}) \sum_{r_2=1}^{d} \frac{\partial V_{i_2,r_2}}{\partial y_{i_2}}(Y_{r_2})\, dX_{r_2}^{K_2}\, dX_1^K$$

$$= \int_s^t \int_s^r \frac{\partial V_{i,r}}{\partial y_{i_2}}(Y_s) \sum_{r_2=1}^{d} \frac{\partial V_{i_2,r_2}}{\partial y_{i_2}}(Y_s)\, dX_{r_2}^{K_2}\, dX_1^K$$

+ Terminos similares.

$$= \sum \text{Coeficientes}(s) \int_s^t \int_s^{r_1} \cdots \int_s^{r_l} dX_{r_{l+1}}^{K_{l+1}} \cdots dX_{r_1}^{K_1}$$

Definimos

$$S_{s,t}^{K_1,\dots,K_l}[X] := \int_s^t \cdots\cdots \int_s^{r_{l-1}} dX_{r_l}^{K_l} \cdots dX_{r_1}^{K_1}$$

Ahora, en lugar de tomar el mapeo:

$$\{X_t\}_{t\geq 0} \xrightarrow[\text{continuo}]{\text{No es}} \{Y_t\}_{t\geq 0}$$

$$\{X_t\}_{t\geq 0} \xrightarrow[\text{agregar info}]{\text{medible}} \left(\{X_t\}_{t\geq 0},\ \text{todas las de } S^{k_1,\dots,k_\ell}_{s,t}[X] \text{ hasta } \ell \leq q.\right)$$

si es continuo.

si q es suficientemente grande (depende de la Hölder continuidad de X

si X es α-Hölder continuo.

$$q = \left\lfloor \frac{1}{\alpha} \right\rfloor$$

$$\{Y_t\}_{t\geq 0}.$$

Aplicaciones en machine learning:

Objeto de interés:

Signaturas $\rightarrow \left\{ S^{k_1,\dots,k_q}_{s,t}[X] \right\}.$ suponer $s=0,\ t=1$

Remark: Signaturas son "características" especiales de $\{X_t\}_{t\geq 0}$.

Problema práctico:

- $Y_t^1 = \alpha + \beta_1^1 Y_{t-1}^1 + \beta_1^2 Y_{t-2}^2 + \varepsilon_t^1 \leftarrow$ clasif. 1

- $Y_t = \alpha + \beta_2^1 Y_{t-1}^1 + \beta_2^2 Y_{t-2}^2 + \varepsilon_t^2 \longleftarrow$ clasif. 2.