


Rough Paths:

- Motivación original ← estudiar "ecuaciones diferenciales" que incluyen perturbaciones aleatorias.
- Motivación reciente ← Realizar aprendizaje de máquina interpretando datos como trayectorias de funciones continuas.

Motivación clásica:

$Y_t = f(t)$ ← función del tiempo.

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

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

$$\Updownarrow$$
$$Y_t - Y_s = \int_s^t v(Y_r) dr$$

Siguiente paso de generalidad:

$$dY_t = V(Y_t) dX_t$$



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

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

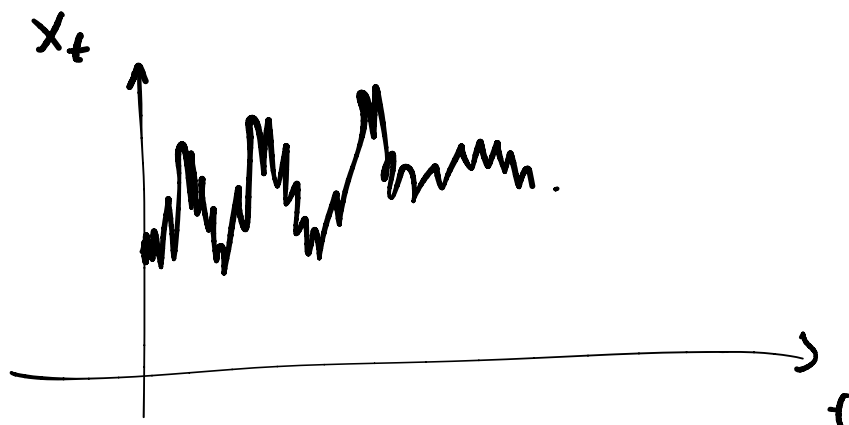
$$dY_t = V(Y_t) dX_t$$

↑ Output ↑ input.

¿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:

$$dY_r = v(Y_r) dX_r \Leftrightarrow Y_{r+\varepsilon} - Y_r \approx v(Y_r) (X_{r+\varepsilon} - X_r).$$

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

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

$$\text{Si } X_i = X_{t_i} \quad Y_i = Y_{t_i}$$

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

$$\Delta X_t \approx \Delta t + \text{error}$$

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

↑
modelo

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

Cuando tomamos límite obtenemos

$$dY_t = v(Y_t) (dt + dB_t).$$

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

es "salt" de definir.

Esto motiva a formular ecuaciones del tipo

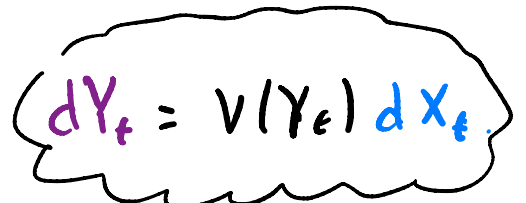
$$dY_t = V(Y_t) dX_t.$$

X_t ← potencialmente tiene trayectorias muy erráticas.

¿Cómo depende el output en función del input?

↕
¿Continuamente?

X


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

Proposición (Lyons et al).

No existe un espacio de Banach \mathcal{B}

f.g. las trayectorias de un mov. Browniano

pertenezcan a \mathcal{B} casi seguramente y f.g.

el mapeo

$$(f, g) \mapsto \int_0^{(\cdot)} f(t)g(t)dt, \text{ definido primero}$$

sobre f, g diferenciables $f, g: \mathbb{R} \rightarrow \mathbb{R}$

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:

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

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

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

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

\leftarrow suponer que tengo acceso a Y_s, X_s .

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

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

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

$$\begin{aligned}
& \int_s^t \int_s^r \frac{\partial V_{i,r}(Y_{r_2})}{\partial y_{i2}} dY_{r_2}^{i_2} dX_r^K \\
&= \int_s^t \int_s^r \frac{\partial V_{i,r}(Y_{r_2})}{\partial y_{i2}} \sum_{r_2=1}^d \frac{\partial V_{i,r_2}(Y_{r_2})}{\partial y_{i2}} dX_{r_2}^{K_2} dX_r^K \\
&= \int_s^t \int_s^r \frac{\partial V_{i,r}(Y_s)}{\partial y_{i2}} \sum_{r_2=1}^d \frac{\partial V_{i,r_2}(Y_s)}{\partial y_{i2}} dX_{r_2}^{K_2} dX_r^K
\end{aligned}$$

+ Terminos similares.

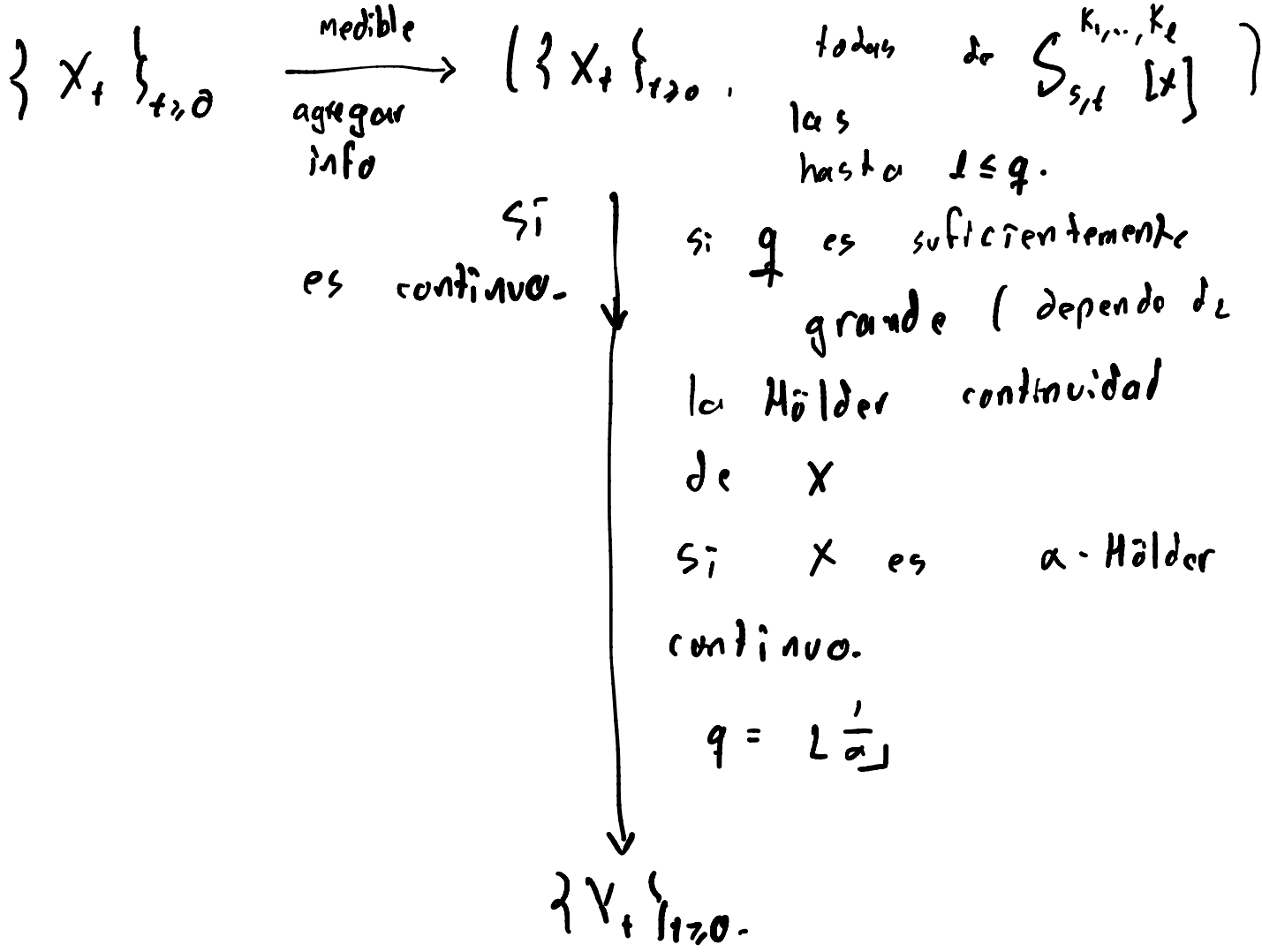
$$= \sum \text{Coeficientes}(s) \int_s^t \int_s^r \dots \int_s^{r_1} dX_{r_{t+1}}^{K_{t+1}} \dots dX_r^{K_t}$$

Definimos

$$\int_{s,t}^{K_1, \dots, K_t} [X] := \int_s^t \dots \int_s^{r_{t-1}} dX_{r_t}^{K_t} \dots dX_r^{K_1}$$

Ahora, en lugar de tomar el mapeo:

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



Aplicaciones en machine learning:

Objeto de interés:

Signaturas $\rightarrow \left\{ S_{s,t}^{K_1, \dots, K_q} [X] \right\}$.

 $\text{Suponer } s=0, t=1$

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

Problema práctico:

- $Y_t^1 = \alpha + \beta_1 Y_{t-1}^1 + \beta_2 Y_{t-2}^1 + \varepsilon_t^1 \leftarrow \text{clasif. 1.}$
 - $Y_t^2 = \alpha + \beta_1 Y_{t-1}^2 + \beta_2 Y_{t-2}^2 + \varepsilon_t^2 \leftarrow \text{clasif. 2.}$
-

Problemas de clasificación que se pueden resolver con técnicas de rough path.

Analogía con regresión:

$Y \leftarrow$ respuesta (score crediticio).

$$Y = \beta_0 + \sum_{j=1}^d \beta_j X_j$$

$X_j \leftarrow$ características,
 $X_j \in \mathbb{R}$.

$\left. \begin{array}{l} Y^1, \dots, Y^n \\ (X_1^1, \dots, X_d^1), \dots, (X_1^n, \dots, X_d^n) \end{array} \right\} \leftarrow \text{observaciones}$

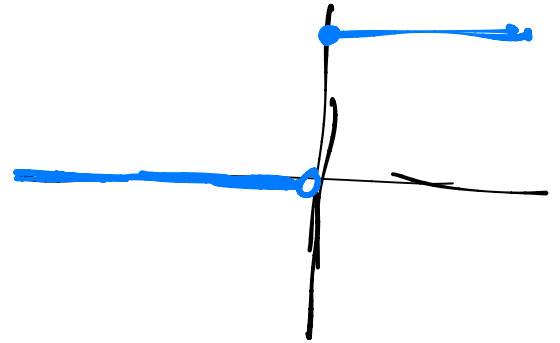
$\beta_0, \dots, \beta_d \leftarrow$ tienen una fórmula para calcularse.

Pasamos ahora de regresión a clasificación

Me gustaría:

$$Y = \phi \left(\beta_0 + \sum_{j=1}^d \beta_j X_j \right)$$

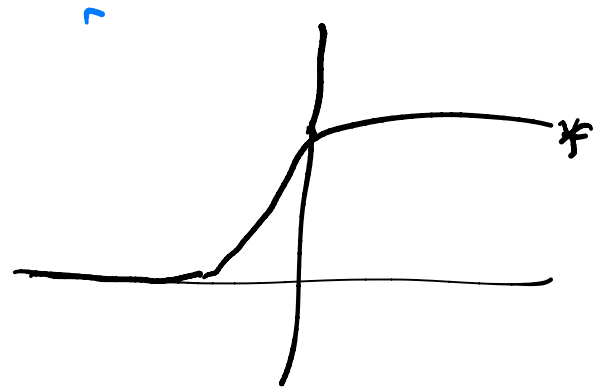
$$\phi(x) = \mathbb{1}_{\{x \geq 0\}}$$



$$Y = \varphi \left(\beta_0 + \sum_{j=1}^d \beta_j X_j \right)$$

$\varphi(x)$ = sigmoide como en

Nos da un aprox. de la clasificación.



Nota:

Si queremos reducir

$$\sum_{l=1}^n \mathcal{E}(Y^l - \varphi(\beta_0 + \sum_{j=1}^d \beta_j X_j^l)) + \sum_{j=1}^d w_j \tilde{\mathcal{E}}(\beta_j)$$

$\mathcal{E} \leftarrow$ cuantificación del error

$\tilde{\mathcal{E}} \leftarrow$ penaliz.

De vuelta al problema de clasificación de series de tiempo:

- $X_t^1 = \alpha + \beta_1 X_{t-1}^1 + \beta_2 X_{t-2}^1 + \varepsilon_t^1 \leftarrow \text{clasif. 1. } \textcircled{\#}$

- $X_t^2 = \alpha + \beta_1 X_{t-1}^2 + \beta_2 X_{t-2}^2 + \varepsilon_t^2 \leftarrow \text{clasif. 2. } \textcircled{\#\#}$

Queremos:

- $Y \leftarrow \text{respuesta: eres } \textcircled{\#} \text{ o } \textcircled{\#\#} \checkmark$

Valores de $Y \leftarrow \{0, 1\} \checkmark$

Ahora vamos a observar:

Voy a ver algo como:



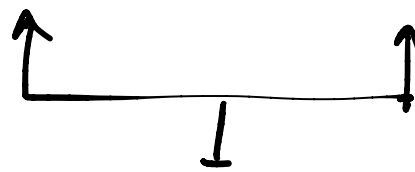
Input \leftarrow la trayectoria de la serie

Problema: como aplico lo que tengo en slides anteriores ?

↓ traducción.

Obtener "características" extraídas de la trayectoria de la serie de tiempo?

$$X \xrightarrow{\tilde{S}} (\tilde{S}_1(X), \dots, \tilde{S}_{40}(X)).$$



características.

Ahora: puedo proponer números β_0, \dots, β_4 t.g.

$$Y = \varphi \left(\beta_0 + \sum_{i=1}^{40} \beta_i \tilde{S}_i(X) \right) + \text{error de medición.}$$

Ésta es "toda" la metodología, siempre y cuando \tilde{S}_i ya estén elegidas,

Conexión con Rough paths:

Comentario:

Hay info. que me gustaría que no influyera en la variable de respuesta.

- Reconocimiento de acciones \leftarrow Marcos de referencia no importan
- Reconocimiento de caracteres chinos. \leftarrow invarianza bajo "dibujar más rápido".

Características: $X \in \mathbb{R}^d$

$$(\tilde{S}_1(x), \dots, \tilde{S}_{40}(x))$$

$$= (S_{i_1, \dots, i_q}(x); 1 \leq q \leq Q \quad i_j \in \{1, \dots, d\})$$

Justificación de que S_{i_j} son "buenos candidatos" \leftarrow depende.

- Ventaja: 1 $S_{i_1, \dots, i_q}(x)$ son invariantes
bajo reparametrizaciones.

- Ventaja 2:

$$(S_{i_1, \dots, i_q} ; q \in \mathbb{N} \setminus \{0\}, i_j \in \{1, \dots, d\}).$$

$$\begin{array}{c} \downarrow \\ \tilde{\mathcal{F}} := \bigoplus_{q=1}^{\infty} (\mathbb{R}^d)^{\otimes q} \quad \leftarrow = \left\{ \sum_{q=1}^{\infty} \sum_{i_1, \dots, i_q=1}^d a_{i_1, \dots, i_q} e_{i_1} \otimes \dots \otimes e_{i_q} \right\} \end{array}$$

Para encajar a las S_{i_1, \dots, i_q} en $\tilde{\mathcal{F}}$

escogemos $a_{i_1, \dots, i_q} = S_{i_1, \dots, i_q}$.

Ventaja:

Si consideramos

$$\int_s^t \int_s^{t_1} \int_s^{\dots} \int_s^{t_q} 1 dX_{t_q} \otimes dX_{t_{q-1}} \otimes \dots \otimes dX_s \in (\mathbb{R}^d)^{\otimes q}$$

$$\Downarrow \\ \equiv \sum_{i_1, \dots, i_q} \left(\int_s^t \int_s^{t_1} \int_s^{\dots} \int_s^{t_q} 1 dX_{t_q}^{i_q} dX_{t_{q-1}}^{i_{q-1}} \dots dX_s^{i_1} \right) e_{i_1} \otimes \dots \otimes e_{i_q}$$

Nota:

El "espacio de Fock" $\tilde{F} = \bigoplus_{q=0}^{\infty} (\mathbb{R}^d)^{\otimes q}$

tiene un producto que le llamamos " \otimes "

$$\begin{array}{ccc} (e_{i_1} \otimes e_{i_2}) \otimes (e_{i_3}) & = & e_{i_1} \otimes e_{i_2} \otimes e_{i_3} \\ \uparrow & & \uparrow \\ \tilde{F} & & \tilde{F} \end{array}$$

Ventaja 2:)

Por inv. bajo reparametrización:

puedo pensar que X_t solo se considera para $t \in [0, 1]$.

$Y_t \leftarrow$ otra "trayectoria" $t \in [1, 2]$ t.g. $Y_1 = X_1$.

Tenemos un producto en las trayectorias.

$$(X \otimes Y)_t = \begin{cases} X_t & \text{si } t \in [0, 1] \\ Y_t & \text{si } t \in [1, 2] \end{cases} \quad t \in [0, 2]$$

↑
producto en
las trayectorias.

$$S(X \otimes Y) = S(X) \otimes S(Y). \leftarrow$$

Signaturas
pueden calcularse
en sub-intervalos
y luego "pegarse"

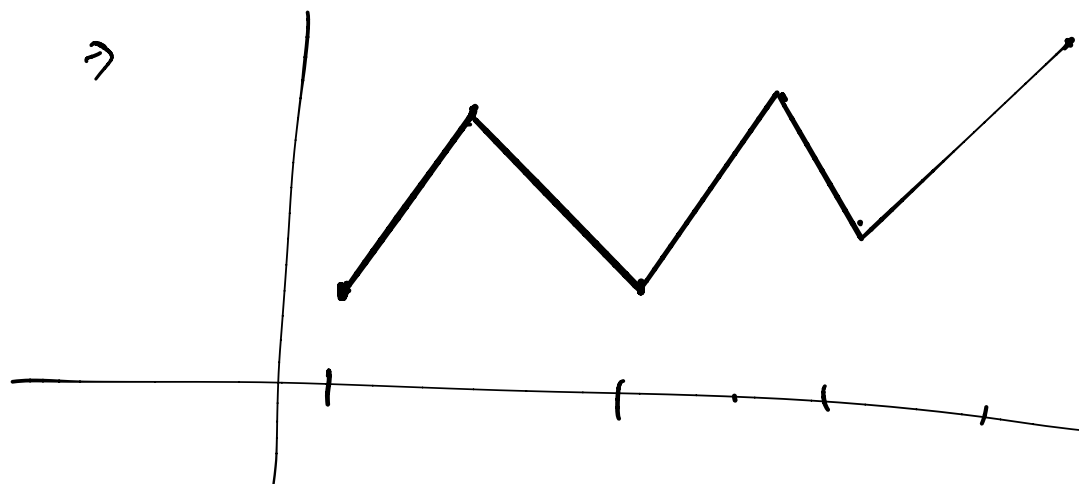
Ventaja 2.2.)

$$X_t = t v \quad v \in \mathbb{R}^d$$

Entonces

$$1 + \int_0^1 dX_v + \int_0^1 \int_0^{t_1} dX_{t_1} \otimes dX_{t_2} + \dots$$

$$= 1 + \sum_{k=1}^M \frac{v^{\otimes k}}{k!} = \exp\{v\}.$$



$X_t \leftarrow$ diferenciable.

$$\int_0^1 X_t dX_t$$

$$S_{i_1, \dots, i_q} S_{j_1, \dots, j_q} = \sum_{\substack{(u_1, \dots, u_q) \\ \in \mathcal{I}}} \square S_{u_1, \dots, u_q, i_1, \dots, i_q}$$

$$C_q = \underset{\text{coef}}{\text{q-ésimo}} \text{ de } \log(E[\exp\{\lambda X_t\}]).$$

Definiciones básicas:

Objetivo:

$X_{(\cdot)}: \mathbb{R}_+ \rightarrow \mathbb{R}$ X_t ← un proceso estocástico.

Para definir la solución de

$$dY_t = V(Y_t) dX_t \iff Y_t - Y_s = \int_s^t V(Y_u) dX_u$$

↑
¿Qué es?

Comentario:

Es muy difícil:

$\int_s^t Z_s dX_s$ de definir para todo Z . \oplus

Nota:

No necesitamos \oplus bien definido para todo Z .
Si estamos considerando el problema

$$Y_t - Y_s = \int_s^t \underbrace{V(Y_u)} dX_u$$

No es cualquier proceso $Z_t = V(Y_t)$.

Claim: será suficiente dar un candidato

$$\text{para } \int_s^t (X_u - X_s) dX_u$$

Heurísticamente, la info que necesitamos para resolver ecuaciones.

$$(X_s, \text{Candidato para } \left(\int_s^t (X_u - X_s) dX_u \right)) = X$$

Al final del seminario, X va a dar el sentido para la ecuación

$$Y_t - Y_s = \int_s^t V(Y_u) dY_u$$

siempre y cuando X no sea muy "rugoso".

Definición:

La rugosidad de una trayectoria la mediremos mediante la Hölder continuidad.

Voy a suponer que $X = \{X_u\}_{u \in [0, T]}$ tiene

Hölder continuidad $\alpha \in (\frac{1}{3}, \frac{1}{2}]$

i.e. $\exists c > 0$ t.q. $\forall s < t < T, |X_t - X_s| \leq c |t - s|^\alpha$

Spoiler:

Si $\alpha \in (\frac{1}{2}, 1)$, podemos definir

$$\int_0^t Y_u dX_u \textcircled{\#} \leftarrow \text{"integral de Young"}$$

si Y y X son α -Hölder continuos.

Spoiler extendido:

Para que $\textcircled{\#}$ exista ocupamos

Y sea p -Hölder continuo

X sea q -Hölder continuo

y $p+q > 1$. En particular, si $p=q=\alpha$.

Definición:

Si V es un espacio de Banach. ($V = \mathbb{R}^d$).

decimos que

$$(X_{(\cdot)}, X_{(\cdot, \cdot)}) : \mathbb{R}_+ \times (\mathbb{R}_+ \times \mathbb{R}_+) \rightarrow V \times (V \otimes V).$$

\uparrow \uparrow
trayect. signatura.
orig.

es un rough path si

$$\cancel{X}_{s,t} - \cancel{X}_{s,u} - \cancel{X}_{u,t} = X_{s,u} \otimes X_{u,t}. \quad \left(\begin{array}{l} \text{Identidad de} \\ \text{Chen} \end{array} \right)$$

donde $X_{s,t} := X_t - X_s$.

Motivación:

Lo que queríamos era $\cancel{X}_{s,t}$ fuera "el" candidato para $\int_s^t X_{s,r} dX_r$

Ejemplo particular del problema:

Si $s \mapsto X_s$ es diferenciable.

$$\cancel{X}_{s,t} - \cancel{X}_{s,u} - \cancel{X}_{u,t}$$

$$\int_s^t X_{s,r} dX_r - \int_s^u X_{s,r} dX_r - \int_u^t X_{u,r} dX_r$$

$$= \int_u^t X_{s,r} dX_r - \int_u^t X_{u,r} dX_r$$

$$= \int_u^t (X_{s,r} - X_{u,r}) dX_r = \textcircled{\#\#}$$

$$\textcircled{II} = \int_u^t (X_r - X_s - \underbrace{X_r + X_u}_{X_{r,u}}) dX_r$$

$$= X_{s,u} \int_u^t dX_r = X_{s,u} X_{u,t}$$

En el caso $V = \mathbb{R}^d$, $d \geq 1$.

$$X_{s,t} \in (\mathbb{R}^d)^{\otimes 2} \approx \text{span} \{ e_i \otimes e_j \mid 1 \leq i, j \leq d \}$$

$$\approx \mathbb{R}^{d \times d} =: \text{Matrices de } d \times d.$$

$$\Rightarrow X_{s,t} = (X_{s,t}^{ij} \mid 1 \leq i, j \leq d).$$

Candidato para $X_{s,t}^{ij}$.

$$X_{s,t}^{ij} = \int_s^t X_{s,r}^i dX_r^j$$

Donde

$$X_t = (X_t^1, \dots, X_t^d).$$

\Rightarrow si definio

$$\int_s^t X_{s,r} \otimes dX_r := \left(\int_s^t X_{s,r}^i dX_r^j \mid 1 \leq i, j \leq d \right).$$

Rough path = (X, \mathbb{X}) , i.g.

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t}. \quad \left(\begin{array}{l} \text{Identidad de} \\ \text{Chen} \end{array} \right)$$

$\mathbb{X}_{s,t}$ ← necesita 2 parâmetros.

No es suficiente conocer

$$\mathbb{X}_{0,t} \quad \forall \quad t > 0$$

X

para conocer a $\mathbb{X}_{s,t}$.

Lo que si es cierto, es que

$(X_t, \mathbb{X}_{0,t})$ determina por completo a

$$(X_{s,t}, \mathbb{X}_{s,t})$$

Razón:

$$X_{s,t} = X_t - X_s \quad \checkmark$$

$$\mathbb{X}_{0,t} - \mathbb{X}_{0,s} - \mathbb{X}_{s,t} = \mathbb{X}_{0,s} \otimes \mathbb{X}_{s,t}$$

$$\Rightarrow \mathbb{X}_{s,t} = \mathbb{X}_{0,t} - \mathbb{X}_{0,s} - \mathbb{X}_{0,s} \otimes \mathbb{X}_{s,t} \leftarrow \begin{array}{l} \text{determinado} \\ \text{por} \\ (X_t, \mathbb{X}_{0,t}). \end{array}$$

Supongamos que (X, \mathbb{X}) es un rough path.

Estoy interpretando $\int_s^t X_{s,r} dY_r = \mathbb{X}_{s,r}$

Claim:

$$\text{Si } Y_{s,t} = \mathbb{X}_{s,t} + F_t - F_s,$$

Para $F_{(\cdot)}: \mathbb{R}_+ \rightarrow V \otimes V$, entonces

(X, Y) satisface la identidad de Chen.

Efectivamente, si

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}.$$

Entonces,

$$Y_{s,t} = \mathbb{X}_{s,t} + F_t - F_s$$

$$Y_{s,u} = \mathbb{X}_{s,u} + F_u - F_s$$

$$Y_{u,t} = \mathbb{X}_{u,t} + F_t - F_u$$

$$\begin{aligned} \Rightarrow Y_{s,t} - Y_{s,u} - Y_{u,t} &= \mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t} \\ &\quad + F_t - F_s - (F_u - F_s + F_t - F_u) \end{aligned}$$

Paso 2:)

Definición:

$\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Definimos el espacio de

α -Hölder rough-paths, escrito como $\mathcal{C}^\alpha([0, T], V)$

como

$(X, \mathbb{X}) \in \mathcal{C}^\alpha([0, T], V)$ si

i) Identidad de Chen se cumple.

y

$$ii). \|X\|_\alpha := \sup_{s < t} \frac{|X_t - X_s|}{|t - s|^\alpha} < \infty.$$

$$\cdot \|\mathbb{X}\|_{2\alpha} := \sup_{s < t} \frac{|\mathbb{X}_{s,t}|}{|t - s|^{2\alpha}} < \infty.$$

$$|X_{s,t}| \lesssim |t - s|^\alpha$$

$$\|\mathbb{X}_{s,t}\| \lesssim |t - s|$$

$$\left| \int_s^t X_{s,r} dX_r \right| \leq \int_s^t \underbrace{|X_{s,r}|}_{\leq |t-s|^\alpha} |dX_r| \leq |t-s|^\alpha |t-s|$$

Si se prueba que *Heurística*

$$\left\| \int_s^t B_{s,t} dB_r \right\|_{L^p(\Omega)} \leq (|t-s|)^{1+p}$$

$t \in [0, T]$.

$$\Rightarrow \sup_{s < t} \frac{\left| \int_s^t B_{s,t} dB_t \right|}{|t-s|^\alpha} < \infty \quad \forall \quad \alpha < \frac{p}{p-1}$$

Pregunta muy difícil de resolver:

$X = \{ X_t \}_{t \geq 0} \in \mathbb{R}^d$ t.q. en $[0, T]$ X tiene trayectorias Hölder continuas de orden $\alpha \in (\frac{1}{3}, \frac{1}{2})$

Pregunta:

Puedo garantizar la existencia de X^* ?

Sí, y un mapeo $X \rightarrow X^*$ que satisfaga

$(X, X^*) \in C^\alpha([0, T], \mathbb{R}^d)$ se llama

levantamiento de Lyons.

En $d=1$

$$X_{s,t}^* = \frac{(X_t - X_s)^2}{2}$$

Identidad de Chen:

$$(X_{s,t}^* - X_{s,u}^* - X_{u,t}^*)$$

$$= \frac{1}{2} \left| (X_t - X_s)^2 - (X_u - X_s)^2 - (X_t - X_u)^2 \right|$$

$$= -\cancel{X_t X_s} - \cancel{X_u^2} + \cancel{X_s X_u} + \cancel{X_t X_u}$$

$$= X_{s,u} \cdot X_{u,t} = (X_u - X_s)(X_t - X_u)$$

Basta ver que para nuestro proceso

X_t se vale

$$\bullet \|X\|_{2\alpha} := \sup_{s < t} \frac{|X_{s,t}|}{|t-s|^{2\alpha}} < \infty.$$

↑
depende del contexto, pero se vale para
mov. Browniano.

$$\int_0^t B_r dB_r \neq \frac{B_t^2}{2}$$