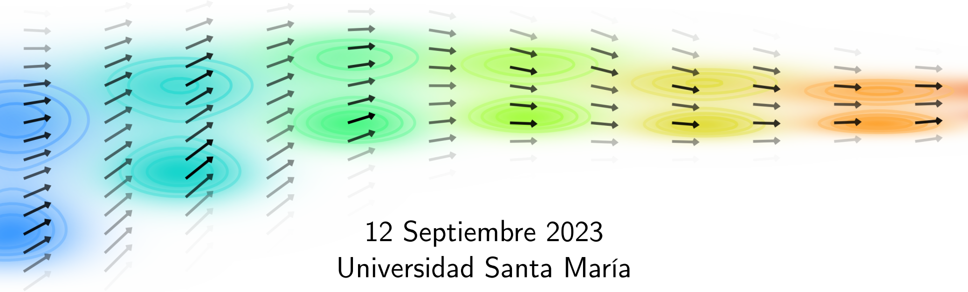
El D en EDP

Estrategías para derivación en espacios de medidas - Capítulo 2, Parte 1

Averil Prost (LMI INSA Rouen)



12 Septiembre 2023
Universidad Santa María

Introducción y notaciones

Notamos $\Omega = \mathbb{R}^d$, $\mathscr{P}_2(\Omega)$ las medidas de probabilidades con segundo momento finito y $d_{\mathcal{W}}(\cdot,\cdot)$ la distancia de Kantorovitch-Rubinstein por $p = 2$, llamada *distancia de Wasserstein* en la literatura.

Introducción y notaciones

Notamos $\Omega = \mathbb{R}^d$, $\mathscr{P}_2(\Omega)$ las medidas de probabilidades con segundo momento finito y $d_{\mathcal{W}}(\cdot, \cdot)$ la distancia de Kantorovitch-Rubinstein por $p = 2$, llamada *distancia de Wasserstein* en la literatura.

La última vez, vimos una definición de gradiente con distribuciones.

Introducción y notaciones

Notamos $\Omega = \mathbb{R}^d$, $\mathscr{P}_2(\Omega)$ las medidas de probabilidades con segundo momento finito y $d_{\mathcal{W}}(\cdot,\cdot)$ la distancia de Kantorovitch-Rubinstein por $p = 2$, llamada *distancia de Wasserstein* en la literatura.

La última vez, vimos una definición de gradiente con distribuciones.

- Resultados muy poderosos, interpretación de ecuaciones como flujos de gradiente de entropías o energías

Introducción y notaciones

Notamos $\Omega = \mathbb{R}^d$, $\mathscr{P}_2(\Omega)$ las medidas de probabilidades con segundo momento finito y $d_{\mathcal{W}}(\cdot,\cdot)$ la distancia de Kantorovitch-Rubinstein por $p = 2$, llamada *distancia de Wasserstein* en la literatura.

La última vez, vimos una definición de gradiente con distribuciones.

- Resultados muy poderosos, interpretación de ecuaciones como flujos de gradiente de entropías o energías
- Difícil de manipular.

Introducción y notaciones

Notamos $\Omega = \mathbb{R}^d$, $\mathscr{P}_2(\Omega)$ las medidas de probabilidades con segundo momento finito y $d_{\mathcal{W}}(\cdot,\cdot)$ la distancia de Kantorovitch-Rubinstein por $p = 2$, llamada *distancia de Wasserstein* en la literatura.

La última vez, vimos una definición de gradiente con distribuciones.

- Resultados muy poderosos, interpretación de ecuaciones como flujos de gradiente de entropías o energías
- Difícil de manipular.

En este capítulo, vamos a descubrir la más conocida de todas las definiciones en el espacio de Wasserstein.

Table of Contents

El teorema fundamental de la simulación

Sea $\mathcal{E}$ una $\sigma-$álgebra y $\mathbb{P} : \mathcal{E} \to \mathbb{R}^+$ una medida.

Def 1 Decimos que $A \in \mathcal{E}$ es un **átomo** de $\mathbb{P}$ si $\mu(A) > 0$ y para cualquier $B \in \mathcal{E}$ tal que $B \subset A$, tenemos $\mu(B) \in \{0, \mu(A)\}$.

El teorema fundamental de la simulación

Sea $\mathcal{E}$ una $\sigma-$álgebra y $\mathbb{P} : \mathcal{E} \to \mathbb{R}^+$ una medida.

Def 1 Decimos que $A \in \mathcal{E}$ es un **átomo** de $\mathbb{P}$ si $\mu(A) > 0$ y para cualquier $B \in \mathcal{E}$ tal que $B \subset A$, tenemos $\mu(B) \in \{0, \mu(A)\}$.

Teorema fundamental de la simulación ([BL94, Theorem A.3.1] o [CD18, Lema 5.29]) Sea $(E, \mathcal{E}, \mathbb{P})$ un espacio medido sin átomo. Para toda $\mu \in \mathscr{P}_2(\Omega)$, existe una variable aleatoria $X \in L^2_{\mathbb{P}}(E; \Omega)$ de ley μ, i.e.

$$\mu = X\#\mathbb{P}, \quad \text{es decir} \quad \mu(O) = \mathbb{P}(X^{-1}(O)) \quad \forall O \subset \Omega \text{ medible.}$$

El teorema fundamental de la simulación

Sea $\mathcal{E}$ una $\sigma-$álgebra y $\mathbb{P} : \mathcal{E} \to \mathbb{R}^+$ una medida.

Def 1 Decimos que $A \in \mathcal{E}$ es un **átomo** de $\mathbb{P}$ si $\mu(A) > 0$ y para cualquier $B \in \mathcal{E}$ tal que $B \subset A$, tenemos $\mu(B) \in \{0, \mu(A)\}$.

Teorema fundamental de la simulación ([BL94, Theorem A.3.1] o [CD18, Lema 5.29]) Sea $(E, \mathcal{E}, \mathbb{P})$ un espacio medido sin átomo. Para toda $\mu \in \mathscr{P}_2(\Omega)$, existe una variable aleatoria $X \in L^2_{\mathbb{P}}(E; \Omega)$ de ley μ, i.e.

$$\mu = X\#\mathbb{P}, \quad \text{es decir} \quad \mu(O) = \mathbb{P}(X^{-1}(O)) \quad \forall O \subset \Omega \text{ medible.}$$

Remark Si $E = [0, 1]$ con borelianos y $\mathbb{P} = \delta_0$, no puede levantar $\mu \neq \delta_x$.

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

Ejemplos

Consideramos $E = [0, 1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$,

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$, podemos elegir cada $X : [0,1] \to \mathbb{R}^d$ $\mathcal{L}-$casi siempre igual a 0.

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$, podemos elegir cada $X : [0,1] \to \mathbb{R}^d$ $\mathcal{L}-$casi siempre igual a 0.
- Si $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$,

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$, podemos elegir cada $X : [0,1] \to \mathbb{R}^d$ $\mathcal{L}-$casi siempre igual a 0.
- Si $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, podemos elegir $X : [0,1] \to \mathbb{R}^d$ con

$$X(s) = \mathbb{1}_{\{s \geqslant 1/2\}} - \mathbb{1}_{\{s<1/2\}}.$$

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$, podemos elegir cada $X : [0,1] \to \mathbb{R}^d$ $\mathcal{L}-$casi siempre igual a 0.
- Si $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, podemos elegir $X : [0,1] \to \mathbb{R}^d$ con

$$X(s) = \mathbb{1}_{\{s \geqslant 1/2\}} - \mathbb{1}_{\{s<1/2\}}.$$

- Si $\Omega = \mathbb{R}$ y $\mu = \rho\mathcal{L}_{\mathbb{R}}$ con $\rho > 0$, podemos elegir

$$X(s) = F^{-1}(s), \qquad F : \mathbb{R} \to [0,1] \text{ dado por } F(r) = \mu\left(]-\infty, r]\right).$$

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$, podemos elegir cada $X : [0,1] \to \mathbb{R}^d$ $\mathcal{L}-$casi siempre igual a 0.
- Si $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, podemos elegir $X : [0,1] \to \mathbb{R}^d$ con

$$X(s) = \mathbb{1}_{\{s \geqslant 1/2\}} - \mathbb{1}_{\{s<1/2\}}.$$

- Si $\Omega = \mathbb{R}$ y $\mu = \rho\mathcal{L}_{\mathbb{R}}$ con $\rho > 0$, podemos elegir

$$X(s) = F^{-1}(s), \qquad F : \mathbb{R} \to [0,1] \text{ dado por } F(r) = \mu\left(] - \infty, r]\right).$$

- Si $\Phi : E \to E$ deja $\mathbb{P}$ invariante, Ley(X) = Ley$(X \circ \Phi)$.

Ejemplos

Consideramos $E = [0,1]$ con la $\sigma-$álgebra de Lebesgue, y $\mathbb{P} = \mathcal{L}_{[0,1]}$.

- Si $\mu = \delta_0 \in \mathscr{P}_2(\Omega)$, podemos elegir cada $X : [0,1] \to \mathbb{R}^d$ $\mathcal{L}-$casi siempre igual a 0.
- Si $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, podemos elegir $X : [0,1] \to \mathbb{R}^d$ con
$$X(s) = \mathbb{1}_{\{s \geqslant 1/2\}} - \mathbb{1}_{\{s < 1/2\}}.$$
- Si $\Omega = \mathbb{R}$ y $\mu = \rho \mathcal{L}_{\mathbb{R}}$ con $\rho > 0$, podemos elegir
$$X(s) = F^{-1}(s), \qquad F : \mathbb{R} \to [0,1] \text{ dado por } F(r) = \mu\left(]-\infty, r]\right).$$
- Si $\Phi : E \to E$ deja $\mathbb{P}$ invariante, Ley$(X) =$ Ley$(X \circ \Phi)$.

¡No hay unicidad!

La L-diferenciabilidad

Def 2 El **lift** de $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es la función

$$U : L^2_{\mathbb{P}}(E;\Omega) \to \mathbb{R}, \qquad U(X) := u\left(\mathsf{Ley}(X)\right).$$

La L-diferenciabilidad

Def 2 El **lift** de $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es la función

$$U : L^2_{\mathbb{P}}(E; \Omega) \to \mathbb{R}, \qquad U(X) := u\left(\text{Ley}(X)\right).$$

Notamos que U es constante en cada conjunto $\text{Ley}^{-1}(\mu)$.

La L-diferenciabilidad

Def 2 El **lift** de $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es la función

$$U : L^2_{\mathbb{P}}(E;\Omega) \to \mathbb{R}, \qquad U(X) := u\left(\mathrm{Ley}(X)\right).$$

Notamos que U es constante en cada conjunto $\mathrm{Ley}^{-1}(\mu)$.

Def 3 Decimos que u es **L-diferenciable** en $\mu \in \mathscr{P}_2(\Omega)$ si existe $X \in \mathrm{Ley}^{-1}(\mu)$ tal que el lift U sea Fréchet-diferenciable en X en el espacio $L^2_{\mathbb{P}}(E;\Omega)$,

La L-diferenciabilidad

Def 2 El **lift** de $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es la función

$$U : L^2_{\mathbb{P}}(E;\Omega) \to \mathbb{R}, \qquad U(X) := u\left(\mathsf{Ley}(X)\right).$$

Notamos que U es constante en cada conjunto $\mathsf{Ley}^{-1}(\mu)$.

Def 3 Decimos que u es **L-diferenciable** en $\mu \in \mathscr{P}_2(\Omega)$ si existe $X \in \mathsf{Ley}^{-1}(\mu)$ tal que el lift U sea Fréchet-diferenciable en X en el espacio $L^2_{\mathbb{P}}(E;\Omega)$, es decir, existe $\nabla_X U \in (L^2_{\mathbb{P}})' \simeq L^2_{\mathbb{P}}(E;T\Omega)$ tal que para cada $H \in \mathcal{C}^1([0,1];L^2_{\mathbb{P}})$,

$$\lim_{s \searrow 0} \frac{U(X+H_s) - U(X)}{|H_s|} = \left\langle \nabla_X U, \dot{H}_0 \right\rangle_{L^2_{\mathbb{P}}}.$$

Ejemplo

Consideramos $\ell \in \mathcal{C}_c^1(\Omega; \mathbb{R})$ y $u(\mu) \coloneqq \int_{x\in\Omega} \ell(x) d\mu(x)$.

Ejemplo

Consideramos $\ell \in \mathcal{C}_c^1(\Omega;\mathbb{R})$ y $u(\mu) \coloneqq \int_{x\in\Omega} \ell(x)d\mu(x)$. Su lift es

$$U(X) = u\left(X\#\mathbb{P}\right) = \int_{z\in E} \ell(X(z))d\mathbb{P}(z).$$

Ejemplo

Consideramos $\ell \in \mathcal{C}_c^1(\Omega;\mathbb{R})$ y $u(\mu) \coloneqq \int_{x\in\Omega} \ell(x)d\mu(x)$. Su lift es

$$U(X) = u\left(X\#\mathbb{P}\right) = \int_{z\in E} \ell(X(z))d\mathbb{P}(z).$$

Luego, para cada $H \in \mathcal{C}^1\left([0,1]; L^2_{\mathbb{P}}(E;\Omega)\right)$ con $H_0 = 0$,

$$\frac{U(X+H_s) - U(X)}{s} = \int_{z\in E} \frac{\ell(X(z)+H_s(z)) - \ell(X(z))}{s} d\mathbb{P}(z)$$

Ejemplo

Consideramos $\ell \in \mathcal{C}_c^1(\Omega;\mathbb{R})$ y $u(\mu) := \int_{x\in\Omega} \ell(x)d\mu(x)$. Su lift es

$$U(X) = u\left(X\#\mathbb{P}\right) = \int_{z\in E} \ell(X(z))d\mathbb{P}(z).$$

Luego, para cada $H \in \mathcal{C}^1\left([0,1]; L^2_{\mathbb{P}}(E;\Omega)\right)$ con $H_0 = 0$,

$$\begin{aligned}\frac{U(X+H_s)-U(X)}{s} &= \int_{z\in E} \frac{\ell(X(z)+H_s(z)) - \ell(X(z))}{s} d\mathbb{P}(z) \\ &\underset{s\searrow 0}{\longrightarrow} \int_{z\in E} \left\langle \nabla_{X(z)}\ell, \dot{H}_0(z)\right\rangle d\mathbb{P}(z).\end{aligned}$$

Ejemplo

Consideramos $\ell \in \mathcal{C}_c^1(\Omega;\mathbb{R})$ y $u(\mu) \coloneqq \int_{x\in\Omega} \ell(x)d\mu(x)$. Su lift es

$$U(X) = u\left(X\#\mathbb{P}\right) = \int_{z\in E} \ell(X(z))d\mathbb{P}(z).$$

Luego, para cada $H \in \mathcal{C}^1\left([0,1]; L^2_{\mathbb{P}}(E;\Omega)\right)$ con $H_0 = 0$,

$$\begin{aligned}\frac{U(X+H_s)-U(X)}{s} &= \int_{z\in E} \frac{\ell(X(z)+H_s(z))-\ell(X(z))}{s} d\mathbb{P}(z)\\ &\xrightarrow[s\searrow 0]{} \int_{z\in E} \left\langle \nabla_{X(z)}\ell, \dot{H}_0(z)\right\rangle d\mathbb{P}(z).\end{aligned}$$

Entonces $\nabla_X U$ es la variable $z \mapsto \nabla_{X(z)}\ell$, que pertenece a $L^2_{\mathbb{P}}(E;T\Omega)$.

Table of Contents

Coherencia

En la Def 3, es suficiente que *existe* $X \in \mathsf{Ley}^{-1}(\mu)$ t.q. U sea dif. en X.

Coherencia

En la Def 3, es suficiente que *existe* $X \in \text{Ley}^{-1}(\mu)$ t.q. U sea dif. en X.

> **[CD18, Proposición 5.24]** Si existe $X \in \text{Ley}^{-1}(\mu)$ tal que U sea diferenciable en X, entonces para todos $Y \in \text{Ley}^{-1}(\mu)$, U es diferenciable en Y.

Coherencia

En la Def 3, es suficiente que *existe* $X \in \text{Ley}^{-1}(\mu)$ t.q. U sea dif. en X.

[CD18, Proposición 5.24] Si existe $X \in \text{Ley}^{-1}(\mu)$ tal que U sea diferenciable en X, entonces para todos $Y \in \text{Ley}^{-1}(\mu)$, U es diferenciable en Y.

Necesitamos la siguiente recíproca parcial de $\tau\#\mathbb{P} = \mathbb{P} \implies X \sim X \circ \tau$.

[CD18, Lema 5.23] Sean $X, Y \in L^2_{\mathbb{P}}(E; \Omega)$ de misma ley.

Coherencia

En la Def 3, es suficiente que *existe* $X \in \text{Ley}^{-1}(\mu)$ t.q. U sea dif. en X.

[CD18, Proposición 5.24] Si existe $X \in \text{Ley}^{-1}(\mu)$ tal que U sea diferenciable en X, entonces para todos $Y \in \text{Ley}^{-1}(\mu)$, U es diferenciable en Y.

Necesitamos la siguiente recíproca parcial de $\tau\#\mathbb{P} = \mathbb{P} \implies X \sim X \circ \tau$.

[CD18, Lema 5.23] Sean $X, Y \in L^2_{\mathbb{P}}(E;\Omega)$ de misma ley. Entonces para todo $\varepsilon > 0$, existe dos aplicaciones medibles $\tau_\varepsilon, \tau_\varepsilon^{-1} : E \to E$ tal que $\mathbb{P}$ sea invariante para τ_ε y τ_ε^{-1}, y

Coherencia

En la Def 3, es suficiente que *existe* $X \in \text{Ley}^{-1}(\mu)$ t.q. U sea dif. en X.

[CD18, Proposición 5.24] Si existe $X \in \text{Ley}^{-1}(\mu)$ tal que U sea diferenciable en X, entonces para todos $Y \in \text{Ley}^{-1}(\mu)$, U es diferenciable en Y.

Necesitamos la siguiente recíproca parcial de $\tau\#\mathbb{P} = \mathbb{P} \implies X \sim X \circ \tau$.

[CD18, Lema 5.23] Sean $X, Y \in L^2_{\mathbb{P}}(E;\Omega)$ de misma ley. Entonces para todo $\varepsilon > 0$, existe dos aplicaciones medibles $\tau_\varepsilon, \tau_\varepsilon^{-1} : E \to E$ tal que $\mathbb{P}$ sea invariante para τ_ε y τ_ε^{-1}, y

$$\mathbb{P} - \mathring{\forall}\, z \in E, \qquad \tau_\varepsilon \circ \tau_\varepsilon^{-1} = \tau_\varepsilon^{-1} \circ \tau_\varepsilon = id,$$

Coherencia

En la Def 3, es suficiente que *existe* $X \in \text{Ley}^{-1}(\mu)$ t.q. U sea dif. en X.

[CD18, Proposición 5.24] Si existe $X \in \text{Ley}^{-1}(\mu)$ tal que U sea diferenciable en X, entonces para todos $Y \in \text{Ley}^{-1}(\mu)$, U es diferenciable en Y.

Necesitamos la siguiente recíproca parcial de $\tau\#\mathbb{P} = \mathbb{P} \implies X \sim X \circ \tau$.

[CD18, Lema 5.23] Sean $X, Y \in L^2_{\mathbb{P}}(E;\Omega)$ de misma ley. Entonces para todo $\varepsilon > 0$, existe dos aplicaciones medibles $\tau_\varepsilon, \tau_\varepsilon^{-1} : E \to E$ tal que $\mathbb{P}$ sea invariante para τ_ε y τ_ε^{-1}, y

$$\mathbb{P} - \mathring{\forall}\, z \in E, \qquad \tau_\varepsilon \circ \tau_\varepsilon^{-1} = \tau_\varepsilon^{-1} \circ \tau_\varepsilon = id, \qquad |X - Y \circ \tau_\varepsilon| \leqslant \varepsilon.$$

Sketch de la demostración

Asumamos que U es Fréchet-diferenciable en $X \in L^2_{\mathbb{P}}(E;\Omega)$, y sea $Y \in \mathsf{Ley}^{-1}(\mathsf{Ley}(X))$.

Sketch de la demostración

Asumamos que U es Fréchet-diferenciable en $X \in L^2_{\mathbb{P}}(E;\Omega)$, y sea $Y \in \text{Ley}^{-1}(\text{Ley}(X))$. Para cada $\varepsilon > 0$, sean $\tau_\varepsilon, \tau_\varepsilon^{-1}$ del lema precedente. Para cada $H \in \mathcal{C}^1([0,1]; L^2_{\mathbb{P}})$ con $H_0 = 0$, tenemos

$$U(Y + H_s) = U(X + Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X)$$

Sketch de la demostración

Asumamos que U es Fréchet-diferenciable en $X \in L^2_{\mathbb{P}}(E;\Omega)$, y sea $Y \in \text{Ley}^{-1}(\text{Ley}(X))$. Para cada $\varepsilon > 0$, sean $\tau_\varepsilon, \tau_\varepsilon^{-1}$ del lema precedente. Para cada $H \in \mathcal{C}^1([0,1]; L^2_{\mathbb{P}})$ con $H_0 = 0$, tenemos

$$\begin{aligned} U(Y + H_s) &= U(X + Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X) \\ &= U(X) + \langle \nabla_X U, Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X \rangle + o\left(\|p\|\right) \end{aligned}$$

donde $p = Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X$ y usamos ambas propriedades de $\tau_\varepsilon, \tau_\varepsilon^{-1}$.

Sketch de la demostración

Asumamos que U es Fréchet-diferenciable en $X \in L^2_{\mathbb{P}}(E;\Omega)$, y sea $Y \in \text{Ley}^{-1}(\text{Ley}(X))$. Para cada $\varepsilon > 0$, sean $\tau_\varepsilon, \tau_\varepsilon^{-1}$ del lema precedente. Para cada $H \in \mathcal{C}^1([0,1];L^2_{\mathbb{P}})$ con $H_0 = 0$, tenemos

$$\begin{aligned} U(Y + H_s) &= U(X + Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X) \\ &= U(X) + \langle \nabla_X U, Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X \rangle + o\left(\|p\|\right) \\ &= U(Y) + \left\langle \nabla_X U \circ \tau_\varepsilon^{-1}, H_s \right\rangle + O(\varepsilon) + o\left(\|H_s\|\right), \end{aligned}$$

donde $p = Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X$ y usamos ambas propriedades de $\tau_\varepsilon, \tau_\varepsilon^{-1}$.

Sketch de la demostración

Asumamos que U es Fréchet-diferenciable en $X \in L^2_{\mathbb{P}}(E;\Omega)$, y sea $Y \in \text{Ley}^{-1}(\text{Ley}(X))$. Para cada $\varepsilon > 0$, sean $\tau_\varepsilon, \tau_\varepsilon^{-1}$ del lema precedente. Para cada $H \in \mathcal{C}^1([0,1]; L^2_{\mathbb{P}})$ con $H_0 = 0$, tenemos

$$\begin{aligned} U(Y + H_s) &= U(X + Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X) \\ &= U(X) + \langle \nabla_X U, Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X \rangle + o\left(\|p\|\right) \\ &= U(Y) + \left\langle \nabla_X U \circ \tau_\varepsilon^{-1}, H_s \right\rangle + O(\varepsilon) + o\left(\|H_s\|\right), \end{aligned}$$

donde $p = Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X$ y usamos ambas propriedades de $\tau_\varepsilon, \tau_\varepsilon^{-1}$.

Es posible probar que $\nabla_X U \circ \tau_\varepsilon^{-1} \underset{\varepsilon \searrow 0}{\longrightarrow} Z$ tal que $(X, \nabla_X U) \sim (Y, Z)$.

Sketch de la demostración

Asumamos que U es Fréchet-diferenciable en $X \in L^2_{\mathbb{P}}(E;\Omega)$, y sea $Y \in \text{Ley}^{-1}(\text{Ley}(X))$. Para cada $\varepsilon > 0$, sean $\tau_\varepsilon, \tau_\varepsilon^{-1}$ del lema precedente. Para cada $H \in \mathcal{C}^1([0,1];L^2_{\mathbb{P}})$ con $H_0 = 0$, tenemos

$$\begin{aligned} U(Y + H_s) &= U(X + Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X) \\ &= U(X) + \langle \nabla_X U, Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X \rangle + o\left(\|p\|\right) \\ &= U(Y) + \left\langle \nabla_X U \circ \tau_\varepsilon^{-1}, H_s \right\rangle + O(\varepsilon) + o\left(\|H_s\|\right), \end{aligned}$$

donde $p = Y \circ \tau_\varepsilon + H_s \circ \tau_\varepsilon - X$ y usamos ambas propriedades de $\tau_\varepsilon, \tau_\varepsilon^{-1}$.

Es posible probar que $\nabla_X U \circ \tau_\varepsilon^{-1} \underset{\varepsilon \searrow 0}{\longrightarrow} Z$ tal que $(X, \nabla_X U) \sim (Y, Z)$.

Pasando al límite en $\varepsilon \searrow 0$, luego en $s \searrow 0$, encontramos

$$\lim_{s \searrow 0} \frac{U(Y + H(s)) - U(Y)}{s} = \left\langle Z, \dot{H}(0) \right\rangle.$$

Estructura del lift

Estructura Sea $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ L-diferenciable en $\mu \in \mathscr{P}_2(\Omega)$.

Estructura del lift

Estructura Sea $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ L-diferenciable en $\mu \in \mathscr{P}_2(\Omega)$. Existe una aplicación $\xi \in L^2_\mu(\Omega; T\Omega)$ tal que

$$\forall X \in \mathsf{Ley}^{-1}(\mu), \qquad \nabla_X U = \xi \circ X.$$

Estructura del lift

Estructura Sea $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ L-diferenciable en $\mu \in \mathscr{P}_2(\Omega)$. Existe una aplicación $\xi \in L^2_\mu(\Omega; T\Omega)$ tal que

$$\forall X \in \mathsf{Ley}^{-1}(\mu), \qquad \nabla_X U = \xi \circ X.$$

Def 4 Llamamos ξ la **L-derivada** de u en μ, denotada por $\partial_\mu u$.

Estructura del lift

Estructura Sea $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ L-diferenciable en $\mu \in \mathscr{P}_2(\Omega)$. Existe una aplicación $\xi \in L^2_\mu(\Omega; T\Omega)$ tal que

$$\forall X \in \mathrm{Ley}^{-1}(\mu), \qquad \nabla_X U = \xi \circ X.$$

Def 4 Llamamos ξ la **L-derivada** de u en μ, denotada por $\partial_\mu u$.

Esa idea aparece en las famosas conferencias de P.L. Lions [Lio06], transcritas en inglés por P. Cardaliaguet [Car13]. Noción muy popular para estudiar juegos de campo medio o ecuaciones de HJ ([CCD15, PW17, PW18, BY19, CGK+22, MZ22]). También existen derivadas de orden más alto [Sal23].

Definición global

El elemento $\xi = \partial_\mu u$ pertenece a $L^2_\mu(\Omega; T\Omega)$, y $\xi(x)$ tiene sentido sólo si $x \in$ suppμ.

Definición global

El elemento $\xi = \partial_\mu u$ pertenece a $L^2_\mu(\Omega; T\Omega)$, y $\xi(x)$ tiene sentido sólo si $x \in$ suppμ. Sin embargo, en caso que u sea más suave, se puede definir $\partial_\mu u(x)$ globalmente de manera natural.

Completamente $\mathcal{C}^1$ Supongamos que $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es globalmente L-diferenciable, que cada $\partial_\mu u$ es continua en supp μ, y que para todos $M \subset \mathscr{P}_2(\Omega)$ y $x \in \bigcap_{\mu \in M}$ suppμ, la función

$$M \ni \mu \mapsto \partial_\mu u(x) \in T\Omega$$

es continua.

Definición global

El elemento $\xi = \partial_\mu u$ pertenece a $L^2_\mu(\Omega; T\Omega)$, y $\xi(x)$ tiene sentido sólo si $x \in$ suppμ. Sin embargo, en caso que u sea más suave, se puede definir $\partial_\mu u(x)$ globalmente de manera natural.

Completamente $\mathcal{C}^1$ Supongamos que $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es globalmente L-diferenciable, que cada $\partial_\mu u$ es continua en supp μ, y que para todos $M \subset \mathscr{P}_2(\Omega)$ y $x \in \bigcap_{\mu \in M}$ suppμ, la función

$$M \ni \mu \mapsto \partial_\mu u(x) \in T\Omega$$

es continua. Entonces existe un representante globalmente continuo de $(\mu, x) \mapsto \partial_\mu u(x)$.

Definición global

El elemento $\xi = \partial_\mu u$ pertenece a $L^2_\mu(\Omega; T\Omega)$, y $\xi(x)$ tiene sentido sólo si $x \in \operatorname{supp}\mu$. Sin embargo, en caso que u sea más suave, se puede definir $\partial_\mu u(x)$ globalmente de manera natural.

Completamente $\mathcal{C}^1$ Supongamos que $u : \mathscr{P}_2(\Omega) \to \mathbb{R}$ es globalmente L-diferenciable, que cada $\partial_\mu u$ es continua en $\operatorname{supp}\mu$, y que para todos $M \subset \mathscr{P}_2(\Omega)$ y $x \in \bigcap_{\mu \in M} \operatorname{supp}\mu$, la función

$$M \ni \mu \mapsto \partial_\mu u(x) \in T\Omega$$

es continua. Entonces existe un representante globalmente continuo de $(\mu, x) \mapsto \partial_\mu u(x)$.

Por ejemplo, es el caso de $u(\mu) = \langle \ell, \mu \rangle$, donde $(\mu, x) \mapsto \partial_\mu u(x) = \nabla_x \ell$.

Regla de la cadena

Asumamos que el lift U es de clase $\mathcal{C}^2$.

Regla de la cadena

Asumamos que el lift U es de clase $\mathcal{C}^2$. Entonces, para todos $(X, Y) \in L^2_{\mathbb{P}}(E; \Omega)$,

$$U(Y) - U(X) = \int_{\tau=0}^{1} \left\langle D_{(1-\tau)X+\tau Y} U, Y - X \right\rangle d\tau.$$

Regla de la cadena

Asumamos que el lift U es de clase $\mathcal{C}^2$. Entonces, para todos $(X, Y) \in L^2_{\mathbb{P}}(E; \Omega)$,

$$U(Y) - U(X) = \int_{\tau=0}^{1} \left\langle D_{(1-\tau)X+\tau Y} U, Y - X \right\rangle d\tau.$$

Para todos $\tau \in [0, 1]$, notamos $\mu_\tau = \text{Ley}\left((1-\tau)X + \tau Y\right)$. Entonces

$$\boxed{u(\mu_1) - u(\mu_0) = \int_{\tau=0}^{1} \left\langle \partial_{\mu_\tau} u \circ \left((1-\tau)X + \tau Y\right), Y - X \right\rangle d\tau.}$$

Regla de la cadena

Asumamos que el lift U es de clase $\mathcal{C}^2$. Entonces, para todos $(X, Y) \in L^2_{\mathbb{P}}(E; \Omega)$,

$$U(Y) - U(X) = \int_{\tau=0}^{1} \left\langle D_{(1-\tau)X+\tau Y} U, Y - X \right\rangle d\tau.$$

Para todos $\tau \in [0, 1]$, notamos $\mu_\tau = \text{Ley}\left((1-\tau)X + \tau Y\right)$. Entonces

$$\boxed{u(\mu_1) - u(\mu_0) = \int_{\tau=0}^{1} \langle \partial_{\mu_\tau} u \circ \left((1-\tau)X + \tau Y\right), Y - X \rangle \, d\tau.}$$

Remark 1 Aquí, $\tau \mapsto \mu_\tau$ depende de la ley conjunta de X y Y. Las trayectorias $\tau \mapsto \mu_\tau$ corresponden a los $\tau \mapsto \left((1-\tau)\pi_x + \tau \pi_y\right) \# \eta$ para $\eta \in \Gamma(\mu, \nu)$: no solamente a lo largo de las geodésicas.

Continuará

Esta parte

- vínculos entre medidas y variables aleatorias
- definición del lift de Lions

Continuará

Esta parte

- vínculos entre medidas y variables aleatorias
- definición del lift de Lions

La definición con el lift es practica y muy utilizada. Sin embargo, se podría esperar una definición más natural, sin recurso a un espacio externo.

Continuará

Esta parte

- vínculos entre medidas y variables aleatorias
- definición del lift de Lions

La definición con el lift es practica y muy utilizada. Sin embargo, se podría esperar una definición más natural, sin recurso a un espacio externo.

La próxima parte

- otra definición del gradiente de funciones suaves

Continuará

Esta parte

- vínculos entre medidas y variables aleatorias
- definición del lift de Lions

La definición con el lift es practica y muy utilizada. Sin embargo, se podría esperar una definición más natural, sin recurso a un espacio externo.

La próxima parte

- otra definición del gradiente de funciones suaves
- vínculos entre los dos

¡Gracias!

[BL94] Nicolas Bouleau and Dominique Lépingle.
Numerical Methods for Stochastic Processes.
1994.

[BY19] Alain Bensoussan and Sheung Chi Phillip Yam.
Control problem on space of random variables and master equation.
ESAIM: Control, Optimisation and Calculus of Variations, 25:10, 2019.

[Car13] Pierre Cardaliaguet.
Notes on Mean Field Games.
page 59, 2013.

[CCD15] Jean-François Chassagneux, Dan Crisan, and François Delarue.
A Probabilistic approach to classical solutions of the master equation for large population equilibria, April 2015.

[CD18] René Carmona and François Delarue.
Probabilistic Theory of Mean Field Games with Applications I, volume 83 of *Probability Theory and Stochastic Modelling*.
Springer International Publishing, Cham, 2018.

[CGK+22] Andrea Cosso, Fausto Gozzi, Idris Kharroubi, Huyên Pham, and Mauro Rosestolato.
Master Bellman equation in the Wasserstein space: Uniqueness of viscosity solutions, February 2022.

[Lio06] Pierre-Louis Lions.
Jeux à champ moyen, 2006.

[MZ22] Chenchen Mou and Jianfeng Zhang.
Wellposedness of Second Order Master Equations for Mean Field Games with Nonsmooth Data, January 2022.

[PW17] Huyên Pham and Xiaoli Wei.
Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics, January 2017.

[PW18] Huyên Pham and Xiaoli Wei.
Bellman equation and viscosity solutions for mean-field stochastic control problem.
ESAIM: Control, Optimisation and Calculus of Variations, 24(1):437–461, January 2018.

[Sal23] William Salkeld.
Higher order Lions-Taylor expansions, March 2023.