Stochastic Processes

MATH5835, P. Del Moral

UNSW, School of Mathematics & Statistics

Lectures Notes, No. 9

Consultations (RC 5112):

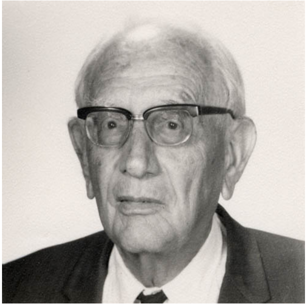
Wednesday 3.30 pm $\rightsquigarrow$ 4.30 pm & Thursday 3.30 pm $\rightsquigarrow$ 4.30 pm

Reminder + Information

References in the slides

- **Material for research projects** $\rightsquigarrow$ Moodle
  (*Stochastic Processes and Applications* $\ni$ variety of applications)
- **Important results**

  $\subset$ **Assessment/Final exam** $=$ LOGO $=$



Mathematics consists in proving the most obvious thing
in the least obvious way. – *George (György) Pólya (1887-1985)*

Plan of the lecture

Purely stochastic techniques

- Coupling distances
  - Total variation distance
  - Wasserstein metric

Plan of the lecture

Purely stochastic techniques

- Coupling distances
  - Total variation distance
  - Wasserstein metric
- Stopping times
  - Coupling times of chains
  - Strong stationary times

Three objectives

- **Choose the right tool** to analyze stability

Three objectives

- **Choose the right tool** to analyze stability
- **Quantify rate of convergence to equilibrium**
  - Coupling distances
  - Find judicious stopping times

Three objectives

- **Choose the right tool** to analyze stability
- **Quantify rate of convergence to equilibrium**
  - Coupling distances
  - Find judicious stopping times
- **develop intuition [without calculations !]**

Coupling

Remember...

$$\mathrm{Law}(X) = \mu_1 \ \& \ \mathrm{Law}(Y) = \mu_2 \ \Rightarrow \ \|\mu_1 - \mu_2\|_{tv} \leq \mathbb{P}(X \neq Y)$$

In fact...

Coupling

Remember...

$$\text{Law}(X) = \mu_1 \ \& \ \text{Law}(Y) = \mu_2 \ \Rightarrow \ \|\mu_1 - \mu_2\|_{tv} \leq \mathbb{P}(X \neq Y)$$

In fact...

Theorem

$$\|\mu_1 - \mu_2\|_{tv} = \inf\{\mathbb{P}(X \neq Y) \ : \ (X, Y) \text{ s.t. } \text{Law}(X) = \mu_1 \ \& \ \text{Law}(Y) = \mu_2\}$$

Proof: Maximal coupling

Blackboard

The Wasserstein metric

Probabilities μ_i on (S, d)

$$\mathbb{W}(\mu_1, \mu_2) = \inf \left\{ \mathbb{E}\left(d(X,Y)\right) \ : \ (X,Y) \text{ s.t. Law}(X) = \mu_1 \ \& \ \text{Law}(Y) = \mu_2 \right\}$$

The Wasserstein metric ⚠

Probabilities μ_i on (S, d)

$$\mathbb{W}(\mu_1, \mu_2) = \inf \{\mathbb{E}\,(d(X,Y)) \;:\; (X,Y) \text{ s.t. Law}(X) = \mu_1 \;\&\; \text{Law}(Y) = \mu_2\}$$

Example:

$$\begin{aligned} \mathbb{W}(\mathcal{N}(m_1,\sigma_1), \mathcal{N}(m_2,\sigma_2)) &= \mathbb{E}\,(|(m_1 - m_2) + (\sigma_1 - \sigma_2)\,\mathcal{N}(0,1))|) \\ &\le |m_1 - m_2| + |\sigma_1 - \sigma_2| \end{aligned}$$

The Wasserstein metric ⚠

Probabilities μ_i on (S, d)

$$\mathbb{W}(\mu_1, \mu_2) = \inf \{\mathbb{E}\,(d(X, Y)) \; : \; (X, Y) \text{ s.t. } \mathrm{Law}(X) = \mu_1 \;\&\; \mathrm{Law}(Y) = \mu_2\}$$

Example:

$$\begin{aligned} \mathbb{W}(\mathcal{N}(m_1, \sigma_1), \mathcal{N}(m_2, \sigma_2)) &= \mathbb{E}\,(|(m_1 - m_2) + (\sigma_1 - \sigma_2)\, \mathcal{N}(0, 1))|) \\ &\leq |m_1 - m_2| + |\sigma_1 - \sigma_2| \end{aligned}$$

Kantorovich-Rubinstein duality theorem:

$$\mathbb{W}(\mu_1, \mu_2) = \sup \{|\mu_1(f) - \mu_2(f)| \; : \; f \in \mathrm{Lip}(S) \text{ s.t. } \mathrm{lip}(f) \leq 1\}$$

Proof $\geq$

The Wasserstein metric

Probabilities μ_i on (S, d)

$$\mathbb{W}(\mu_1, \mu_2) = \inf \{\mathbb{E}\,(d(X, Y)) \ : \ (X, Y) \text{ s.t. } \mathrm{Law}(X) = \mu_1 \ \& \ \mathrm{Law}(Y) = \mu_2\}$$

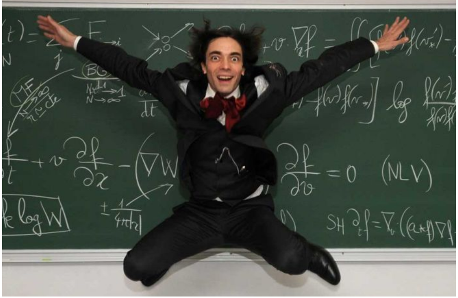
Example:

$$\begin{aligned} \mathbb{W}(\mathcal{N}(m_1, \sigma_1), \mathcal{N}(m_2, \sigma_2)) &= \mathbb{E}\,(|(m_1 - m_2) + (\sigma_1 - \sigma_2)\, \mathcal{N}(0, 1))|) \\ &\leq |m_1 - m_2| + |\sigma_1 - \sigma_2| \end{aligned}$$

Kantorovich-Rubinstein duality theorem:

$$\mathbb{W}(\mu_1, \mu_2) = \sup \{|\mu_1(f) - \mu_2(f)| \ : \ f \in \mathrm{Lip}(S) \text{ s.t. } \mathrm{lip}(f) \leq 1\}$$

Proof $\geq$ ($\leq$ cf. C. Villani book!) Blackboard

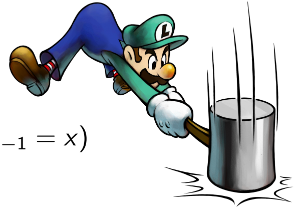


Kantorovich-Rubinstein duality theorem:

$$\mathbb{W}(\mu_1, \mu_2) = \sup \{ |\mu_1(f) - \mu_2(f)| \ : \ f \in \mathrm{Lip}(S) \text{ s.t. } \mathrm{lip}(f) \leq 1 \}$$

Connexion with optimal transport : cf. C. Villani book!)

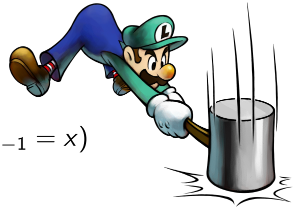
Intermediate tool



Prop.:

$$(\delta_x M)(dy) = M(x, dy) = \mathbb{P}(X_n \in dy \mid X_{n-1} = x)$$

Intermediate tool



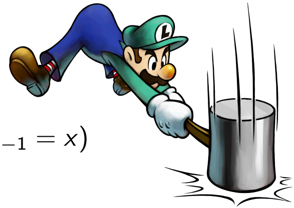
Prop.:

$$(\delta_x M)(dy) = M(x, dy) = \mathbb{P}(X_n \in dy \mid X_{n-1} = x)$$

$$\Downarrow$$

$$\mathbb{W}(\delta_x M, \delta_y M) \leq a\ d(x, y) \Longrightarrow \mathbb{W}(\mu_1 M, \mu_2 M) \leq a\ \mathbb{W}(\mu_1, \mu_2)$$

Intermediate tool



Prop.:

$$(\delta_x M)(dy) = M(x, dy) = \mathbb{P}(X_n \in dy \mid X_{n-1} = x)$$

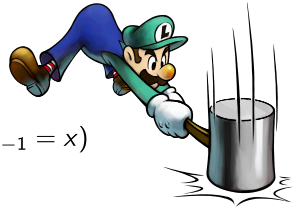
$$\Downarrow$$

$$\mathbb{W}(\delta_x M, \delta_y M) \leq a\ d(x, y) \Longrightarrow \mathbb{W}(\mu_1 M, \mu_2 M) \leq a\ \mathbb{W}(\mu_1, \mu_2)$$

$$\Downarrow$$

$$\mathbb{W}(\mu_0 M^n, \mu_0' M^n) \leq a^n\ \mathbb{W}(\mu_0, \mu_0')$$

Intermediate tool



Prop.:

$$(\delta_x M)(dy) = M(x, dy) = \mathbb{P}(X_n \in dy \mid X_{n-1} = x)$$

$$\Downarrow$$

$$\mathbb{W}(\delta_x M, \delta_y M) \leq a \; d(x,y) \Longrightarrow \mathbb{W}(\mu_1 M, \mu_2 M) \leq a \; \mathbb{W}(\mu_1, \mu_2)$$

$$\Downarrow$$

$$\mathbb{W}(\mu_0 M^n, \mu_0' M^n) \leq a^n \; \mathbb{W}(\mu_0, \mu_0')$$

Example $U \sim \gamma(du)$:

$$X_n = F(X_{n-1}, U_n) \quad \text{with} \quad \int \|F(x,u) - F(y,u)\| \, \gamma(du) \leq a \; \|x - y\|$$

Proof: Blackboard

Stopping times and coupling

Stopping time: $\{T = n\} \sim (X_0, \ldots, X_n)$

(As far as we know, photo is public domain)

Stopping times and coupling ⚠

(As far as we know, photo is public domain)

Stopping time: $\{T = n\} \sim (X_0, \ldots, X_n)$

$$\Downarrow$$

Coupling time T of two chains

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv} \leq \mathbb{P}(X_n \neq Y_n) \leq \mathbb{P}(T > n) \leq \mathbb{E}(T)/n$$

Stopping times and coupling

(As far as we know, photo is public domain)

Stopping time: $\{T = n\} \sim (X_0, \ldots, X_n)$

$$\Downarrow$$

Coupling time T of two chains

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv} \leq \mathbb{P}\left(X_n \neq Y_n\right) \leq \mathbb{P}\left(T > n\right) \leq \mathbb{E}(T)/n$$

Ex.: 2 copies of the same chain

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv}$$

$$\leq \int \mathbb{P}(X_0 \in dx)\mathbb{P}(Y_0 \in dy)\, \|\mathrm{Law}(X_n \mid X_0 = x) - \mathrm{Law}(Y_n \mid Y_0 = y))\|_{tv}$$

Stopping times and coupling

(As far as we know, photo is public domain)

Stopping time: $\{T = n\} \sim (X_0, \ldots, X_n)$

$$\Downarrow$$

Coupling time T of two chains

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv} \leq \mathbb{P}(X_n \neq Y_n) \leq \mathbb{P}(T > n) \leq \mathbb{E}(T)/n$$

Ex.: 2 copies of the same chain

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv}$$

$$\leq \int \mathbb{P}(X_0 \in dx)\mathbb{P}(Y_0 \in dy)\,\|\mathrm{Law}(X_n \mid X_0 = x) - \mathrm{Law}(Y_n \mid Y_0 = y))\|_{tv}$$

$$\leq \sup_{(x,y)\in S^2} \mathbb{P}(T_{x,y} > n)$$

Stopping times and coupling

(As far as we know, photo is public domain)

Stopping time: $\{T = n\} \sim (X_0, \ldots, X_n)$

$$\Downarrow$$

Coupling time T of two chains

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv} \leq \mathbb{P}\left(X_n \neq Y_n\right) \leq \mathbb{P}\left(T > n\right) \leq \mathbb{E}(T)/n$$

Ex.: 2 copies of the same chain

$$\|\mathrm{Law}(X_n) - \mathrm{Law}(Y_n)\|_{tv}$$

$$\leq \int \mathbb{P}(X_0 \in dx)\mathbb{P}(Y_0 \in dy) \, \|\mathrm{Law}(X_n \mid X_0 = x) - \mathrm{Law}(Y_n \mid Y_0 = y))\|_{tv}$$

$$\leq \sup_{(x,y)\in S^2} \mathbb{P}\left(T_{x,y} > n\right)$$

Examples: when $M(x, dy) \geq \epsilon \, \lambda(dy) \ldots$ (cf. previous lectures)

Strong stationary times

X_n **with invariant measure** π**:**

$$T \text{ strong stationary times} \iff X_T \perp T \quad \& \quad \text{Law}(X_T) = \pi$$

Strong stationary times

X_n with invariant measure π:

$$T \text{ strong stationary times} \iff X_T \perp T \quad \& \quad \text{Law}(X_T) = \pi$$

Ex.: top-in-card shuffle $T = 1+$ the first time top card back to top

Strong stationary times

X_n with invariant measure π:

$$T \text{ strong stationary times} \iff X_T \perp T \quad \& \quad \mathrm{Law}(X_T) = \pi$$

Ex.: top-in-card shuffle $T = 1+$ the first time top card back to top

Prop.:

$$\|\mathrm{Law}(X_n) - \pi\|_{tv} \leq \mathbb{P}(T > n)$$

Proof: Blackboard