Bayesian Learning for Partially-Observed Dynamical Systems

Randal DOUC and Sylvain Le Corff

Tutorial 4 : Asymptotic properties of Markov chains.

randal.douc@telecom-sudparis.eu sylvain.le_corff@telecom-sudparis.eu

CHAPITRE 4 : ASYMPTOTIC PROPERTIES OF MARKOV CHAINS

For simplicity, we assume that $\mathsf{X} = \mathbb{R}^k$.

Définition 0.1 Let P be a Markov kernel on $\mathsf{X} \times \mathcal{B}(\mathsf{X})$. We say that $\bar{P}$ is a *coupling kernel* of (P,P) if and only if

$$\forall (x,x') \in \mathsf{X}^2, \forall A \in \mathcal{B}(\mathsf{X}) \quad \begin{cases} \bar{P}((x,x'), A \times \mathsf{X}) = P(x,A) \\ \bar{P}((x,x'), \mathsf{X} \times A) = P(x',A) \end{cases}$$

EXERCICE 1 Let ξ, ξ' two probability measures on the measurable set $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$. Let ν be a probability measure on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$ such that $\mathrm{d}\xi = \varphi \mathrm{d}\nu$ and $\mathrm{d}\xi' = \varphi' \mathrm{d}\nu$, that is φ (resp. φ') is the density of ξ (resp. ξ') wrt ν. The total variation distance between ξ and ξ' is defined by :

$$\|\xi - \xi'\|_{TV} = \int |\varphi(x) - \varphi'(x)| \nu(\mathrm{d}x) = \sup_{|h| \le 1} |\xi(h) - \xi'(h)| = 2 \inf_{\mu \in \mathcal{C}(\xi,\xi')} \int \mathbb{1}(x \neq x') \mu(\mathrm{d}x \mathrm{d}x')$$

where $\mathcal{C}$ is the set of coupling distributions μ of (ξ, ξ'), that is, distributions μ on $(\mathsf{X}^2, \mathcal{B}(\mathsf{X})^{\otimes 2})$ such that the following marginal conditions are satisfied : for all $A \in \mathcal{B}(\mathsf{X})$,

$$\mu(A \times \mathsf{X}) = \xi(A)$$
$$\mu(\mathsf{X} \times A) = \xi'(A)$$

1. Show the different equalities that appear in the definition of the Total Variation distance.

EXERCICE 2 Let P be a Markov kernel on $\mathsf{X} \times \mathcal{B}(\mathsf{X})$ such that

(i) There exist a measurable function $V : \mathsf{X} \to [1, \infty)$, $\lambda \in (0,1)$ and $b \in \mathbb{R}$ such that for all $x \in \mathsf{X}$,

$$PV(x) \le \lambda V(x) + b$$

(ii) There exist $M > 0$, $\varepsilon > 0$ and a probability measure ν on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$ such that for all $x \in C_M \stackrel{\text{def}}{=} \{x \in \mathsf{X} : V(x) \le M\}$,

$$P(x, \cdot) \ge \varepsilon \nu(\cdot)$$

(iii) $\bar{\lambda} \stackrel{\text{def}}{=} \lambda + \frac{2b}{1+M} < 1$.

In what follows, we use the notation $\bar{x} = (x,x')$ and $\bar{C}_M = C_M \times C_M$. Writing for all $x \in C_M$, $Q(x, \mathrm{d}y) = \frac{P(x,\mathrm{d}y) - \varepsilon \nu(\mathrm{d}y)}{1-\varepsilon}$, we define the Markov kernel $\bar{P}$ on $\mathsf{X}^2 \times \mathcal{B}(\mathsf{X})^{\otimes 2}$ by

$$\bar{P}((x,x'); \mathrm{d}y\mathrm{d}y') = \mathbb{1}_{\bar{C}_M}(\bar{x}) \left[\varepsilon \nu(\mathrm{d}y) \delta_y(\mathrm{d}y') + (1-\varepsilon) Q(x,\mathrm{d}y) Q(x',\mathrm{d}y')\right] + \mathbb{1}_{\bar{C}_M^c}(\bar{x}) P(x,\mathrm{d}y) P(x',\mathrm{d}y')$$

1. Show that $\bar{P}$ is a kernel coupling of (P, P).
2. Define $\mathrm{d}(x,x') = \mathbb{1}(x \neq x')$. Show that $\bar{P}\mathrm{d}(x,x') \leq (1-\varepsilon)\mathrm{d}(x,x')$ for all $\bar{x} = (x,x') \in C_M$.
3. Define $\bar{V}(x,x') = \frac{V(x)+V(x')}{2}$. Show that $\bar{P}\bar{V}(x,x') \leq \bar{\lambda}\bar{V}(x,x')$ for all $\bar{x} = (x,x') \notin C_M$.
4. Deduce that there exists $\delta, \rho \in (0,1)$ such that defining $W = V^{\delta}\mathrm{d}^{1-\delta}$, we have

$$\bar{P}W \leq \rho W$$

5. Show that for all $x, x' \in \mathsf{X}$ and all $n \in \mathbb{N}$,

$$\|P^n(x,\cdot) - P^n(x',\cdot)\|_{TV} \leq 2\rho^n W(x,x')$$

6. Deduce that P admits an invariant probability measure π. **We admit that the set of probability measures on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$ equipped with the Total Variation distance is a complete space.**
7. Show that $\pi(V) < \infty$ and deduce that for all $x \in \mathsf{X}$ and all $n \in \mathbb{N}$,

$$\|P^n(x,\cdot) - \pi\|_{TV} \leq \rho^n \{V(x) + \pi(V)\}$$

EXERCICE 3 In this exercise, we consider the same assumptions as in Exercise 2. For any real-valued function h on X such that $\pi(|h|) < \infty$, we say that a real-valued function $\hat{h}$ on X solves the Poisson equation associated to f if for all $x \in \mathsf{X}$,

$$\hat{h}(x) - P\hat{h}(x) = h(x) - \pi(h)$$

provided that $P\hat{h}(x)$ is well-defined for all $x \in \mathsf{X}$. Define $S_n(h) = \sum_{k=0}^{n-1}\{h(X_k) - \pi(h)\}$.

1. Show that

$$S_n(h) = M_n(h) + \hat{h}(X_0) - \hat{h}(X_n)$$

where $M_n(h) = \sum_{k=1}^{n}\left\{\hat{h}(X_k) - P\hat{h}(X_{k-1})\right\}$.
2. Show that $\{M_n(h)\}_{n\geq 0}$ is an $(\mathcal{F}_n)_{n\geq 0}$-martingale, where we have set $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.
3. Show that for all bounded function h, the function $\hat{h}(x) = \sum_{k=0}^{\infty}\left\{P^k(h)(x) - \pi(h)\right\}$ solves the Poisson equation.