Chap 1-2 : Introduction aux CM.

Dans tout le cours : $(\mathcal{X}, \mathfrak{X})$ espace mesurable.

- $\mathcal{M}_+(\mathfrak{X})$: ensemble des mesures (+) sur $(\mathcal{X}, \mathfrak{X})$
- $\mathcal{M}_1(\mathfrak{X})$: ———— de proba ————
- $F_+(\mathfrak{X})$: ———— f^{ns} à valeurs réelles ⊕, $\mathfrak{X}/\mathcal{B}(\mathbb{R}^+)$ mesurables.
- $F_b(\mathfrak{X})$: ———— ————, $\mathfrak{X}/\mathcal{B}(\mathbb{R})$ mesurables, bornées.

Def : Une fonction : $P : \mathcal{X} \times \mathfrak{X} \longrightarrow \mathbb{R}^+$ est un noyau de Markov
$(x, A) \longmapsto P(x,A)$
ssi : $\forall (x,A) \in \mathcal{X} \times \mathfrak{X}$,
$$\begin{cases} y \mapsto P(y,A) & \text{est } \mathfrak{X}/\mathcal{B}(\mathbb{R}^+) \text{ mesurable.} \\ B \mapsto P(x,B) & \text{est une proba} \end{cases}$$

Rappel notation : $\forall \mu \in \mathcal{M}_+(\mathcal{X})$, $\mu(dx)$ est la notation infinitésimale pour écrire : $\mu(f) = \int f(x)\mu(dx)$.
De m̂ : $P(x,A) = \int \mathbb{1}_A(y)\, P(x, dy)$.

- $\forall \nu \in \mathcal{M}_1(\mathcal{X})$, $\begin{cases} P \text{ noyau de Markov sur } \mathcal{X} \times \mathfrak{X}, \quad h \in F_+(\mathfrak{X}). \\ Q \end{cases}$

Alors : μP est la mesure : $A \mapsto \mu P(A) = \int \mu(dx)\, P(x,A)$
PQ est le noyau : $(x,A) \mapsto PQ(x,A) = \int P(x,dy)\, Q(y,A)$.
Qh est la f^{n} mesurable : $x \longmapsto Qh(x) = \int Q(x,dy)\, h(y)$.

Autre notation : $\forall k \geq 1$, $P^k = P \circ \dots \circ P$. avec $P^0 = Id$ ie : $P^0(x,dy) = \delta_x(dy)$.

Déf : Une suite de v.a $(X_k)_{k \in \mathbb{N}}$ définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $\mathcal{X}$ forment une CM de noyau P, de loi initiale μ ssi :
(i) $\forall n \geq 1$, $\forall h \in F_+(\mathfrak{X})$, $\mathbb{E}(h(X_n) \mid X_0, \dots, X_{n-1}) = P(X_{n-1}, h) = Ph(X_{n-1})$
(ii) $\forall h \in F_+(\mathfrak{X})$, $\mathbb{E}(h(X_0)) = \int h(x_0)\, \mu(dx_0)$.

Rq : Si $\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}$ filtration sur $(\mathcal{X}, \mathfrak{X})$ (ie : $\forall n \geq 0$, $\mathcal{F}_n$ tribu sur $\mathcal{X}$ et $\mathcal{F}_n \subset \mathcal{F}_{n+1}$).
tq : (X_n) est $\mathcal{F}$-adapté, alors : si $\begin{cases} \forall h \in F_+(X), \\ \forall n \geq 1 \end{cases}$, $\mathbb{E}(h(X_n) \mid \mathcal{F}_{n-1}) = Ph(X_{n-1})$. (*).

on a $\mathbb{E}(h(X_n) \mid X_{0:n-1}) = Ph(X_{n-1})$.

(Preuve : $\mathbb{E}(h(X_n) \mid X_{0:n-1}) = \mathbb{E}\Big(\mathbb{E}(h(X_n) \mid \underbrace{\sigma(X_{0:n-1})}_{\subset \mathcal{F}_{n-1}} \cup \mathcal{F}_{n-1}) \mid X_{0:n-1} \Big)$.
$= \mathbb{E}(Ph(X_{n-1}) \mid X_{0:n-1}) = Ph(X_{n-1}).)$.

Donc pour mq (X_k) est une CM, il suffit de trouver une filtration $\mathcal{F}$ tq : (X_k) $\mathcal{F}$-adapté et (*).

Lemme : Si (X_n) est une chaîne de Markov sur $(\Omega, \mathcal{F}, \mathbb{P})$ prenant valeurs sur $\mathcal{X}$, de noyau P et loi initiale $\nu \in \mathcal{M}_1(\mathcal{X})$. Alors pour tout $n \geq 0$, la loi de $X_{0:n}$ est

$$A \longmapsto \int_A \nu(dx_0) \prod_{i=0}^{n-1} P(x_i, x_{i+1})$$

Proof: Posons H_n : la loi de $X_{0:n}$ est $A \mapsto \int_A \nu(dx_0) \prod_{i=0}^{n-1} P(x_i, dx_{i+1})$

- H_0 vrai (clair)
- supposons H_{n-1} vrai $(n \geq 1)$.

Pour tout $h_0, \dots, h_n \in F_+(\mathcal{X})$, $\mathbb{E}\left[\prod_{i=0}^{n} h_i(X_i)\right] = \mathbb{E}\left[\mathbb{E}\left[\prod_{i=0}^{n} h_i(X_i) \mid X_{0:n-1}\right]\right]$.

$$= \mathbb{E}\Big[\prod_{i=0}^{n-1} h_i(X_i) \underbrace{\mathbb{E}(h_n(X_n) \mid X_{0:n-1})}_{Ph_n(X_{n-1})}\Big) \underset{\text{par } H_{n-1}}{=} \int \nu(dx_0) \prod_{i=0}^{n-2} P(x_i, dx_{i+1}) h_i(x_i) \times h_{n-1}(x_{n-1}) \underbrace{Ph_n(x_{n-1})}_{\int P(x_{n-1}, dx_n) h_n(x_n)}.$$

$$= \int \nu(dx_0) \prod_{i=0}^{n-1} P(x_i, dx_{i+1}) h_i(x_i)) \, h_n(x_n).$$ ce qui montre (H_n)

Déf : Soit $\pi \in M_+(\mathcal{X})$, π est une mesure invariante pour P ssi $\pi P = \pi$.
Si (X_n) chaîne de M. de noyau P et $\pi P = \pi$ alors : si $X_0 \sim \pi$, on a : $X_k \sim \pi$ $\forall k \in \mathbb{N}$.

Lemme : Si P admet 2 probas invariantes, il admet aussi 2 probas invariantes étrangères π_0, π_1.
(i.e : $\begin{cases} \pi_0 P = \pi_0 \\ \pi_1 P = \pi_1 \end{cases}$ et $\exists A \in \mathcal{X}$ s.t : $\pi_0(A) = 0$, $\pi_1(A^c) = 0$.).

Preuve :
Soit $\nu_0, \nu_1 \in M_1(\mathcal{X})$, $\begin{cases} \nu_0 P = \nu_0 \\ \nu_1 P = \nu_1 \end{cases}$ et $\nu_0 \neq \nu_1$. Posons $\nu = \nu_0 + \nu_1$. On a : $\begin{cases} \nu_0 \ll \nu \\ \nu_1 \ll \nu \end{cases}$ donc $\exists f_1, f_2$ tq :
$d\nu_0 = f_0 \, d\nu$
$d\nu_1 = f_1 \, d\nu$.

posons : $(\nu_0 - \nu_1)^+$ la mesure : $A \mapsto \int_A (f_0 - f_1)^+ d\nu$.

alors : $\forall A \in \mathcal{X}$, $(\nu_0 - \nu_1)^+ P(A) \geq (\nu_0 - \nu_1)^+ P(A \cap \{f_0 \geq f_1\})$.

$$\geq (\nu_0 - \nu_1) P(A \cap \{f_0 \geq f_1\}) = (\nu_0 - \nu_1)(A \cap \{f_0 \geq f_1\})$$

$$= \int \Big[\underbrace{(f_0 - f_1) \mathbb{1}_{\{f_0 \geq f_1\}}}_{(f_0 - f_1)^+} \mathbb{1}_A\Big](x) \, \nu(dx).$$

$$= (\nu_0 - \nu_1)^+(A). \quad (1).$$

En appliquant à A^c : $\underbrace{(\nu_0 - \nu_1)^+ P(A^c)}_{} \geq (\nu_0 - \nu_1)^+(A^c)$.

$$\underbrace{(\nu_0 - \nu_1)^+ P(\mathcal{X})}_{\int (\nu_0 - \nu_1)^+(dx) \underbrace{P(x, \mathcal{X})}_{=1}} - (\nu_0 - \nu_1)^+ P(A) \geq (\nu_0 - \nu_1)^+(\mathcal{X}) - (\nu_0 - \nu_1)^+(A).$$

Donc : $\cancel{(\nu_0-\nu_1)^+}(\mathcal{X}) - (\nu_0-\nu_1)^+P\,(A) \geq \cancel{(\nu_0-\nu_1)^+}(\mathcal{X}) - (\nu_0-\nu_1)^+(A)$.

$\underline{ie}$: $(\nu_0-\nu_1)^+(A) \geq (\nu_0-\nu_1)^+P(A)$.

En combinant avec (1), on tire : $(\nu_0-\nu_1)^+P = (\nu_0-\nu_1)^+$

En posant $\pi_0 = \dfrac{(\nu_0-\nu_1)^+}{(\nu_0-\nu_1)^+(\mathcal{X})}$, on a : $\pi_0 \in M_1(\mathcal{X})$ et $\pi_0 P = \pi_0$.

De même, $\pi_1 = \dfrac{(\nu_1-\nu_0)^+}{(\nu_1-\nu_0)^+(\mathcal{X})}$ on a : $\pi_1 \in M_1(\mathcal{X})$ et $\pi_1 P = \pi_1$.

De $\oplus$, $\pi_0(\underbrace{f_0<f_1}_{A}) = \dfrac{(\nu_0-\nu_1)^+(\{f_0<f_1\})^{=0}}{(\nu_0-\nu_1)^+(\mathcal{X})} = 0$. De même, $\pi_1(\underbrace{f_0\geq f_1}_{A^c}) = \dfrac{(\nu_1-\nu_0)^+(f_0\geq f_1)^{=0}}{(\nu_1-\nu_0)^+(\mathcal{X})} = 0$

ce qui achève la preuve.

$\underline{\text{Déf}}$: $\pi \in M_1(\mathcal{X})$ est P-réversible ssi : $\int_A \pi(dx)\,P(x,dy) = \int_A \pi(dy)\,P(y,dx)$, $\forall A \in \mathcal{X}^{\otimes 2}$.

($\underline{ie}$: $\pi(dx)\,P(x,dy) = \pi(dy)\,P(y,dx)$ en notation infinitésimale).

$\underline{\text{Lemme}}$: π est P-réversible $\Rightarrow$ π est P-invariant.

$\underline{\text{Proof}}$: $\forall A \in \mathcal{X}$, $\pi P(A) = \iint \pi(dx)\,P(x,dy)\,\mathbb{1}_A(y) = \iint \pi(dy)P(y,dx)\,\mathbb{1}_A(y)$.

$$= \int \pi(dy) \underbrace{\left[\int P(y,dx)\right]}_{\underbrace{P(y,\mathcal{X})}_{=1}} \mathbb{1}_A(y).$$

$$= \pi(A).$$

$\underline{\text{Exemple}}$: $\underline{\text{MH}}$ algorithm. Si $\pi \in M_1(\mathcal{X})$ et Q noyau sur $\mathcal{X} \times \mathcal{X}$.

On suppose : $\exists \lambda \in M_+(\mathcal{X})$ tq : $\begin{cases} \pi(dx) = \pi(x)\,\lambda(dx). \\ Q(x,dy) = q(x,y)\lambda(dy). \end{cases}$ On suppose : $\forall x \in \mathcal{X}$, $q(x,\cdot) > 0$, λ-pp.

$\underline{\text{Algo}}$: $X_0 \sim \mu$.

Pour $i = 1, \dots, n$.

. tirer $Y_n \sim Q(X_{n-1},\cdot)$ et $U_n \sim \text{Unif}[0,1]$ indép.

. Si $\begin{cases} U_n \leq \alpha(X_{n-1},Y_n) & \text{poser } X_n = Y_n \\ U_n > \alpha(X_{n-1},Y_n) & \text{— } X_n = X_{n-1} \end{cases}$ où $\alpha(x,y) = \dfrac{\pi(y)q(y,x)}{\pi(x)q(x,y)} \wedge 1$.

1) Mq : (X_n) CM de noyau P où $\boxed{P(x,dy) = Q(x,dy)\alpha(x,y) + \bar\alpha(x)\,\delta_x(dy)}$. (**)

où $\bar\alpha(x) = 1 - \int Q(x,dz)\alpha(x,z)$.

2) Mq : P est π-réversible.

Preuve : 1) $\forall h \in F_+(\mathcal{X})$,

$$\mathbb{E}(h(X_n) | X_{0:n-1}) = \mathbb{E}(h(X_n)|X_{n-1}) = \mathbb{E}\Big(h(\underset{Y_n}{\underbrace{X_n}})\mathbb{1}_{U_n \le \alpha(X_{n-1},Y_n)} \Big| X_{n-1}\Big) + \mathbb{E}\Big(h(\underset{X_{n-1}}{\underbrace{X_n}})\mathbb{1}_{U_n > \alpha(X_{n-1},Y_n)} \Big| X_{n-1}\Big)$$

$$= \int Q(X_{n-1}, dy_n)\,\alpha(X_{n-1},y_n)\,h(y_n) + h(X_{n-1}).\underbrace{\underbrace{\int Q(X_{n-1},dy_n)(1-\alpha(X_{n-1},y_n))}_{1 - \int Q(X_{n-1},dz_n)\alpha(X_{n-1},z_n)}}_{\bar\alpha(X_{n-1})}.$$

$$= \int h(y)\Big[\underbrace{Q(X_{n-1},dy)\,\alpha(X_{n-1},y) + \bar\alpha(X_{n-1})\,\delta_{X_{n-1}}(dy)}_{P(X_{n-1},dy)}\Big].$$

2) On remarque que : $\pi(dx)\,Q(x,dy)\,\alpha(x,y) = \pi(dy)\,Q(y,dx)\,\alpha(y,x)$. (a) (Eq. de balance détaillée)

(en effet : $\pi(x)\,q(x,y)\,\alpha(x,y) = \min\big(\pi(y)q(y,x), \pi(x)q(x,y)\big)$, symétrique).

De ⊕, $\pi(dx)\,\bar\alpha(x)\,\delta_x(dy) = \pi(dy)\,\bar\alpha(y)\,\delta_y(dx)$. (b).

(en effet : $\forall f \in F_+(\mathcal{X}^2)$, $\iint f(x,y)\,\pi(dx)\,\bar\alpha(x)\,\delta_x(dy)$

$= \int f(x,x)\,\pi(dx)\,\bar\alpha(x)$.

$= \int f(y,y)\,\pi(dy)\,\bar\alpha(y)$.

$= \int f(x,y)\,\pi(dy)\,\bar\alpha(y)\,\delta_y(dx)$).

Par (a) et (b), et (**), on tire que P est π-réversible (donc π-invariante).