II Rappels sur les chaînes de Markov (Bref : cf cours 4et+) (= Processus Markovien)

1) Définition et propriété :

déf : Un processus aléatoire (= suite de v.a [adapté à une filtration]) est une chaîne de Markov

à valeur dans $(E, \mathcal{E})$ ssi $\forall A \in \mathcal{E}$ $\quad \mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} \in A \mid X_n)$ $\mathbb{P}$-ps

où $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$

→ Propriété de Markov : $\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} \in A \mid X_n)$ $\mathbb{P}$-ps

2) Matrice / Noyau de transition

- **déf :** Lorsque X est discret, $P : \mathcal{X} \times \mathcal{X}$ est une matrice de transition si $\forall (x, y) \in \mathcal{X}^2$

$$P(x,y) \geq 0 \quad \text{et} \quad \sum_{y \in \mathcal{X}} P(x,y) = 1$$

En particulier $\forall x \in \mathcal{X}$ $\quad Q(y) = P(x,y)$ est une loi de $\mathbb{P}$ sur $\mathcal{X}$.

- Lorsque x n'est pas discret, on étend cette définition comme suit

déf : $P : \mathcal{X} \times \mathscr{X}$ est un noyau de transition si

(i) $\forall x \in \mathcal{X}$, $A \mapsto P(x, A)$ est une loi de $\mathbb{P}$ sur $(\mathcal{X}, \mathscr{X})$

(ii) $\forall A \in \mathscr{X}$, $x \mapsto P(x, A)$ est une application mesurable.

ex : Si $\mathcal{X} = \mathbb{R}^d$ et $\mathscr{X}$ = Boréliens de $\mathbb{R}^d$ et si $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ est mesurable tq

$0 < \int f(x,y)\, dy < \infty$ alors $P(x, A) = \dfrac{\int_A f(x,y)\, dy}{\int_{\mathcal{X}} f(x,y)\, dy}$ est un noyau de transition.

Propriété : Action sur les fonctions et mesures : $*$ $\forall f : \mathcal{X} \to \mathbb{R}^+$, on définit une application mesurable $Pf : \mathcal{X} \to \mathbb{R}^+$ par $Pf(x) = \sum_{y \in \mathcal{X}} P(x,y) f(y)$ ou $Pf(x) = \int_{\mathcal{X}} P(x, dy) f(y)$

→ s'étend à f de signe qcq

$*$ $\forall \mu$ mesure sur $(\mathcal{X}, \mathscr{X})$ de $\mathbb{P}$ on définit une mesure de $\mathbb{P}$ μP sur $(\mathcal{X}, \mathscr{X})$ par $\mu P(y) = \sum_{x \in \mathcal{X}} P(x,y) \mu(x)$

ou $\mu P(A) = \int_{\mathcal{X}} \mu(dx) P(x, A)$

- Composition / Itération : Soient (P_1, P_2) deux noyaux (resp matrices) de transition.

On définit

$$[P_1 P_2](x,y) = \sum_{z \in \mathcal{X}} P_1(x,z) P_2(z,y)$$

$$[P_1 P_2](x,A) = \int P_1(x, dz) P_2(z, A)$$

et on peut vérifier que c'est un noyau / matrice de transition

Rq: Dans le cas discret P_1P_2 est un produit matriciel (éventuellement infini), de même que Pf et μP.

Puis par récurrence, la composition de n matrice / noyau de transition est notée P^n:

(⚠ $P_1P_2 \neq P_2P_1$) $\quad P^0(x,A) = \mathbb{1}_A(x) = \delta_x(A)$

$$P^n(x,A) = [P P^{n-1}](x,A) = [P^{n-1}P](x,A)$$

et on a la relation de Chapman - Kolmogorov $\quad P^{n+m}(x,A) = \int P^n(x\, dz)\, P^m(z,A)$

3) Chaîne homogène :

Soit une chaîne de Markov dans $(X, \mathcal{X})$. On a alors la propriété de Markov

$$\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} \in A \mid X_n) = F(n; X_n, A)$$

def: Lorsque $F(n, X_n, A) = F(X_n, A)$ la chaîne est dite homogène. Une chaîne de Markov homogène et de noyau de transition P et de loi initiale μ si

(i) $\mathbb{P}(X_0 \in A) = \mu(A) \quad \forall A \in \mathcal{X}$

(ii) $\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbf{P}(X_n, A) \quad \forall A \in \mathcal{X}$ et $n \geq 0$

Prop : Soit (X_n) une chaîne de Markov de loi initiale μ et de noyau P. Soit $n \geq 0$ un entier et f une fonction mesurable positive sur X^{n+1}. Alors

$$\mathbb{E}[f(X_0, \dots, X_n)] = \int_{X^{n+1}} \mu(dx)\, P(x, dx_1)\, P(x_1, dx_2) \dots P(x_{n-1}, dx_n)\, f(x_0, \dots, x_n)$$

En particulier $\mathbb{E}[f(X_n)] = [\mu P^n](f)$. Donc la loi de X_n est μP^n (dans le cas homogène)

Rq: En discret $\int \rightarrow \sum_{(x_0, \dots, x_n) \in X^{n+1}} \mu(dx_0)\, P(x_0, x_1) \dots P(x_{n-1}, x_n)\, f(x_0, \dots, x_n)$

Preuve : En utilisant les outils standards de théorie de la mesure, il suffit de montrer ceci pour les $f(x_0, \dots, x_n) = \prod_{k=0}^{n} f_k(x_k)$

→ On montre ceci par récurrence sur n : qd $n = 0$ c'est la définition de X_0 et qd $n > 0$ on utilise la formule de conditionnement (formule des probas totales) + noyau de transition + HRecurrence.

4) Mesure invariante :

def: Soit P un noyau de transition sur $(X, \mathcal{X})$ et π une mesure sur $(X, \mathcal{X})$. π est dite invariante pour P ssi $\pi P = \pi$ ie $\forall x \in X$, $\sum_{x \in X} \pi(x) P(x,y) = \pi(y)$

ou $\forall A \in \mathcal{X} \quad \int_X \pi(dx)\, P(x, A) = \pi(A)$

def : Un processus est dit stationnaire si $\forall h \in \mathbb{N}$, $\forall p \in \mathbb{N}$ $(X_{t_1}, \dots X_{t_p})$ a même loi que $(X_{t_1+h}, \dots, X_{t_p+h})$

En particulier, dans le cas d'une chaîne de Markov de loi initiale μ et de noyau de transition P, lorsque le processus est stationnaire, X_1 et X_0 ont même loi donc $\mu P = \mu$ (par la prop précédente) et donc μ est la mesure invariante.

Réciproquement, si $\mu P = \mu$ pour une loi de $\mathbb{P}$ μ alors $\forall n$ $\mu P^n = \mu$ et toutes les v.a X_n ont même loi. Par la prop, on montre de plus que le processus est stationnaire.

Donc une chaîne de Markov est stationnaire ssi sa loi initiale est invariante sous P

Rq : Ceci paraît un peu contradictoire. On veut simuler selon π et si la chaîne est stationnaire de loi invariante π il faut simuler x_0 selon π !

En pratique, on choisit x_0 arbitrairement. C'est la propriété d'ergodicité (de nouveau) qui va nous dire qu'à partir d'un moment (temps à déterminer en fonction du problème) la chaîne est dans un état stationnaire. On a oublié d'où on était parti et maintenant on simule selon π et π est donc bien la loi (stationnaire) invariante de P.

def : Ergodicité : Supposons qu'il existe une mesure de $\mathbb{P}$ π tq $\lim_{n\infty} \sup_{A \in \mathcal{X}} |\mu P^n(A) - \pi(A)| = 0$ (A priori π dépend de μ). Nous verrez (cours 4) que cela est équivalent à $\sup_{f, \|f\|_\infty \leq 1} |\mu P^n f - \pi(f)| \xrightarrow[n\infty]{} 0$.

→ Alors $\forall A \in \mathcal{X}$: $\pi(A) = \lim_{n\infty} \mu P^{n+1}(A) = \lim_{n\infty} \int \mu P^n(dx)\, P(x, A) = \int \pi(dx)\, P(x, A)$

$= \pi P(A)$ donc π est mesure invariante de P.

. Lorsque la mesure limite π est indépendante de μ, la chaîne est dite ergodique. L'ergodicité signifie donc que la chaîne oublie sa loi initiale et que la loi de la chaîne converge vers π. La proba limite est nécessairement invariante pour P.

Dans une optique de "méthodes de Monte Carlo", l'ergodicité est cruciale : elle justifie que pour n grand, les réalisations de X_n sont "à peu près" des réalisations sous π (dont la simulation exacte est impossible)

Question : Peut-on construire une chaîne de Markov ergodique dont la mesure stationnaire est

→ C'est l'objet des échantillonneurs MCMC

→ Autre question : A quelles conditions un noyau P possède-t-il une mesure de probabilité invariante donnée ?

5) Réversibilité et Equation de balance détaillée.

def : Un processus est réversible si $\forall (t_1, \dots, t_p)$ et $\forall T \geq t_p$, $(X_{t_1}, \dots, X_{t_p})$ a même loi que $(X_{T-t_1}, \dots, X_{T-t_p})$

prop : Un processus réversible est stationnaire :

Preuve : Soient $t_1 < \dots < t_p$ et $h \geq 0$. Si le processus est réversible alors $\forall T \geq t_p$ et $T' \geq t_p +$

$(X_{t_1}, \dots, X_{t_p})$ a même loi que $(X_{T-t_1}, \dots, X_{T-t_p})$; De plus $(X_{t_1+h}, \dots, X_{t_p+h})$

a même loi que $(X_{T'-t_1-h}, \dots, X_{T'-t_p-h})$

Pour $T'-h = T \geq t_p$ on obtient la formulation de la stationnarité.

Rq : [Interprétation de la réversibilité dans le cas discret] la définition est équivalente

à dire que $\forall n$ $(X_0, \dots, X_n)$ a même loi que $(X_n, \dots, X_0)$ donc

$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_0 = x_n, \dots, X_n = x_0)$

la réversibilité traduit que le fait de parcourir un chemin dans un sens ou dans l'autre a la même probabilité : le temps est réversible

Cor : Si la chaîne est réversible, alors (X_0, X_1) a même loi que (X_1, X_0) ce qui entraîne que

$\forall (i,j) \in \mathbb{X}$ $\mu(i) P(i,j) = \mu(j) P(j,i)$ $(*)$

Réciproquement, si $(*)$ est vérifiée, on montre par récurrence sur n que $(X_0, \dots X_n)$ a même loi que $(X_n, \dots, X_0)$ et donc que la chaîne est réversible

Rq :
- $(*)$ est connue sous le nom de "équation de balance détaillée". C'est une condition suffisante pour établir la réversibilité et donc la stationnarité d'une chaîne à espace d'états discrets.
- Si on $\sum_i (*)$ on obtient $\mu P = \mu$ et donc μ est invariante pour P.

  C'est donc une condition suffisante pour établir qu'une mesure est invariante pour P.
- $(*)$ exprime un état d'équilibre ("balance") du système : quand le système est à l'équilibre le flux de j vers i est égal au flux de i vers j.