TD 02 : Chaînes de Markov

Propriétés de décorrélation des chaînes de Markov

1. **CHAÎNES DE MARKOV**

Soit Σ un ensemble fini de cardinalité au moins 2. Une *matrice stochastique* sur Σ est une matrice carrée, dont les lignes et les colonnes sont indicées par Σ, à coefficients positifs, et dont la somme de chaque ligne vaut 1.

Une matrice carrée à coefficients positifs A est dite *irréductible* si pour tous i, j dans Σ, il existe $n \geq 1$ tel que $(A^n)_{ij} > 0$, et *apériodique* si pour tous i, j dans Σ, il existe $n_0 \geq 1$ tel que $(A^n)_{ij} > 0$ pour tout $n \geq n_0$. Le théorème de Perron-Forbenius relie ces propriétés au spectre de la matrice.

Théorème 1 (Théorème de Perron-Frobenius)**.**
Soit A une matrice carrée à coefficients positifs et irréductible. Soit ρ son rayon spectral. Alors $\rho > 0$ est une valeur propre simple de A. De plus, il existe un (co)vecteur propre associé à coordonnées strictement positives. Tout vecteur propre associé à une autre valeur propre a des coordonnées dont les phases sont distinctes. Enfin, il existe un entier $p \geq 1$ tel que toute valeur propre de module ρ est simple, et est égale à ρ multiplié par une racine p-ième de l'unité.

De plus, une matrice irréductible est apériodique si et seulement si ρ est l'unique valeur propre de module ρ.

Une *chaîne de Markov* sur Σ de matrice de transition A est un processus stochastique $(M_n)_{n\geq 0}$ sur Σ tel que, pour tout cylindre $[c_0, \ldots, c_{n-1}]$,

$$\mathbb{P}(M_0 = c_0, \ldots, M_{n-1} = c_{n-1}) = \mathbb{P}(M_0 = c_0) \prod_{k=0}^{n-2} A_{c_k, c_{k+1}}. \tag{0.1}$$

(a) Considérons une chaîne de Markov irréductible. Montrer que $\rho(A) = 1$.
(b) En utilisant le théorème de Perron-Frobenius, montrer qu'il existe une unique mesure de probabilité μ sur Σ qui soit invariante au sens suivant : si M_0 est de loi μ, alors M_1 l'est aussi.

On construit alors une mesure $\overline{\mu}$ sur $\Omega := \Sigma^{\mathbb{N}}$ en imposant que M_0 suive la loi μ, et en définissant la mesure des cylindres grâce à l'équation (0.1). Soit T le décalage sur Ω.

(c) Montrer que $(\Omega, \overline{\mu}, T)$ est un système dynamique préservant la mesure. *Indication: on montrera que $\overline{\mu}(T^{-1}A) = \overline{\mu}(A)$ pour tout cylindre A.*
(d) Soient f et g deux fonctions complexes sur Σ. Exprimer $\mathbb{E}(f \cdot g \circ T^n)$ à l'aide de μ et A^n. Si A est apériodique, que peut-on dire quand n tend vers l'infini ?

Soit $\mathcal{L}$ l'opérateur de transfert, défini par:

$$\int_\Omega f \cdot g \circ T \,\mathrm{d}\overline{\mu} = \int_\Omega \mathcal{L}(f) \cdot g \,\mathrm{d}\overline{\mu} \quad \forall f \in \mathbb{L}^1(\Omega, \overline{\mu}), \quad \forall g \in \mathbb{L}^\infty(\Omega, \overline{\mu}.$$

Pour tout $N \geq 0$, on note $\mathcal{F}_N$ l'espace des fonctions complexes sur Ω qui ne dépendent que des N premières coordonnées.

(e) Montrer que $\mathcal{L}$ préserve $\mathcal{F}_1$.
(f) On suppose A apériodique. Montrer que, pour tout $N \geq 0$, il existe des constantes $C \geq 0$ et $r \in [0, 1)$ telles que, pour toutes fonctions $f \in \mathcal{F}_N$ et $g \in \mathbb{L}^1(\Omega, \overline{\mu})$, pour tout $n \geq 0$,

$$|Cov(f, g \circ T^n)| := |\mathbb{E}(f \cdot g \circ T^n) - \mathbb{E}(f)\mathbb{E}(g)| \leq Cr^n \|f\|_{\mathbb{L}^\infty} \|g\|_{\mathbb{L}^1}.$$

On pourra tout d'abord montrer ces inégalités pour des fonctions indicatrices de cylindres.
(g) En déduire que si A est apériodique, alors $(\Omega, \overline{\mu}, T)$ est mélangeant.
(h) Soient $z \in \mathbb{S}_1$ et $f \in \mathbb{L}^2(\Omega, \overline{\mu})$. Montrer que $\mathcal{L}(f) = zf$ si et seulement si $f \circ T = \overline{z}f$. *Indication : pour le sens direct, on pourra calculer $\|f \circ T - \overline{z}f\|_{\mathbb{L}^2}$.*
(i) On suppose maintenant que A est seulement irréductible, et on considère la valeur propre $e^{i2\pi k/p}$. Montrer que l'opérateur de Koopman $f \mapsto f \circ T$ admet une fonction propre associée à $e^{-i2\pi k/p}$, telle que $f^p \equiv 1$.
(j) En déduire qu'il existe une partition $(\Sigma_k)_{k\in\mathbb{Z}_{/p\mathbb{Z}}}$ de Σ telle que T envoie chaque $[\Sigma_k]$ sur $[\Sigma_{k+1}]$. Que peut-on dire de A^p ?
(k) En déduire que si A est irréductible, alors $(\Omega, \overline{\mu}, T)$ est ergodique.
(l) Que vaut l'entropie $h_{\overline{\mu}}(T)$?

2. **UN THÉORÈME CENTRAL LIMITE**

Le but de cet exercice est de démontrer un théorème central limite pour des chaînes de Markov[1]. Soit Σ un ensemble fini, de cardinal au moins 2. Soit A la matrice de transition d'une chaîne de Markov apériodique sur A. Soit μ la mesure de probabilité stationnaire de la chaîne de Markov, c'est-à-dire l'unique covecteur dont la somme des coordonnées vaut 1 et tel que $\mu A = \mu$. On note $D(\mu)$ la matrice diagonale telle que, pour $s \in \Sigma$, on ait $D(\mu)_{ss} = \mu(s)$.

Soit F une fonction réelle sur Σ telle que $\mathbb{E}_\mu(F) = 0$. Soit $\mathcal{F}$ l'espace des fonctions complexes sur Σ. Pour tout $\omega \in \mathbb{R}$, on définit des opérateurs sur $\mathcal{F}$ par :

$$\begin{aligned} \mathcal{L}(g) &= D(\mu)^{-1} A^t (D(\mu) g) \ ; \\ \mathcal{L}_\omega(g) &= \mathcal{L}(e^{i\omega F} \cdot g). \end{aligned}$$

(a) Montrer que 1 est valeur propre simple de $\mathcal{L}$, et que toute autre valeur propre est de module strictement inférieur.

(b) Montrer que, pour tous f et g dans $\mathcal{F}$,

$$\mathbb{E}_{\overline{\mu}}(f \cdot \mathcal{L}(g)) = \mathbb{E}_{\overline{\mu}}(f \circ T \cdot g).$$

En déduire que, pour tous $n \geq 0$ et $\omega \in \mathbb{R}$, pour tout $g \in \mathcal{F}$,

$$\mathcal{L}_\omega^n(g) = \mathcal{L}^n \left(e^{i\omega S_n F} g \right),$$

et que $\mathbb{E}_{\overline{\mu}}(e^{i\omega S_n F}) = \mathbb{E}_{\overline{\mu}}(\mathcal{L}_\omega^n(1))$.

On admettra que, pour tout ω suffisamment petit, il existe une application $\omega \mapsto \lambda(\omega) \in \mathbb{C}$ analytique, telle que $\lambda(\omega)$ soit la valeur propre de $\mathcal{L}_\omega$ de plus grand module. On admettra de plus qu'il existe un application $\omega \mapsto g(\omega) \in \mathcal{F}$ elle aussi analytique, telle que $g(\omega)$ soit l'unique fonction propre associée à $\lambda(\omega)$ telle que $\mathbb{E}_{\overline{\mu}}(g(\omega)) = 1$.

Soit P_ω la projection sur le sous-espace propre associé à $\lambda(\omega)$, et $Q_\omega := \mathcal{L}_\omega - \lambda(\omega) P_\omega$. Alors $\omega \mapsto P_\omega$ est analytique sur un voisinage de 0. En particulier, il existe $C > 0$ et $\theta \in [0, 1)$ tels que $\|Q_\omega^n\| \leq C\theta^n$ pour tout $n \geq 0$ et tout ω assez petit.

(c) Que valent $\lambda(0)$ et $g(0)$?

(d) Montrer que, pour tout ω assez petit, $\mathcal{L}_\omega^n(\mathbf{1}) = \lambda(\omega)^n (\mathbf{1} + O(\omega)) + O(\theta^n)$. Dans la suite, on cherche à développer $\lambda(\omega)$ près de 0.

(e) Montrer que $\mathcal{L}$ agit sur $\mathcal{F}_0 := \{f \in \mathcal{F} \ : \ \mathbb{E}_{\overline{\mu}}(f) = 0\}$, et que $I - \mathcal{L}$ est inversible sur cet espace.

(f) On rappelle que :

$$\lambda(\omega) g(\omega) = \mathcal{L}_\omega g(\omega). \tag{0.2}$$

En dérivant cette équation et en prenant l'espérance, montrer que $\lambda'(0) = 0$. Exploiter ensuite l'équation (0.2) pour montrer que $g'(0) = i(I - \mathcal{L})^{-1} \mathcal{L}(F)$.

(g) Développer l'équation (0.2) au second ordre et prendre l'espérance. En déduire que :

$$\sigma(F)^2 := -\lambda''(0) = \mathbb{E}_\mu(F^2) + 2 \sum_{n \geq 1} \mathbb{E}_\mu(F \cdot \mathcal{L}^n F).$$

(h) En utilisant la question (d), montrer que $\mathbb{E}_{\overline{\mu}} \left(e^{i\omega \frac{S_n F}{\sqrt{n}}} \right)$ converge vers $e^{-\frac{\sigma(F)^2 \omega^2}{2}}$ quand n tend vers $+\infty$. Conclure.

Codage symbolique

3. **APPLICATIONS AFFINES PAR MORCEAUX**

Soit (a, b) un intervalle réel et Σ un ensemble fini. Soit $(a_s)_{s \in \Sigma}$ une partition de (a, b) en sous-ensemble (a_s, b_s). Soit T : $(a, b) \to (a, b)$ une application. On suppose que[2] :

- T est affine sur (a_s, b_s) pour tout $s \in \Sigma$;
- Pour tout $s \in \Sigma$, il existe un ensemble fini E tel que l'ensemble $T((a_s, b_s) \setminus E)$ est une union finie d'éléments de la partition $(a_s)_{s \in \Sigma}$.

Pour tout $x \in [a, b)$, on note $\sigma(x)$ l'unique élément $s \in \Sigma$ tel que $x \in [a_s, b_s)$. On note enfin $\mathcal{F}_\sigma$ l'espace des fonctions de $[a, b)$ dans $\mathbb{R}_+$, constantes sur chaque intervalle $[a_s, b_s)$, et d'intégrale 1.

(a) Montrer que l'opérateur de transfert agit sur $\mathcal{F}_\sigma$, et que cette action a un point fixe.

[1]Ce théorème peut se voir comme donnant une vitesse de convergence dans le théorème de Birkhoff. On verra plus tard un phénomène très différent concernant les rotations irrationelles sur le cercle.

[2]En pratique, on rajoute souvent une hypothèse d'expansivité, qui assure que le codage construit par la suite est essentiellement bijectif. Ce n'est pas important pour cet exercice.

(b) En déduire qu'il existe au moins une mesure de probabilité T-invariante μ, absolument continue par rapport à la mesure de Lebesgue, et dont la densité est constante sur chaque intervalle $[a_s, b_s)$.

(c) Soit $n \geq 0$ et $(s_0, \ldots, s_{n-1}) \in \Sigma^n$. Supposons que $[s_0, \ldots, s_{n-1}] := \bigcap_{k=0}^{n-1} T^{-k}([a_{s_k}, b_{s_k}))$ est de μ-mesure non nulle. Que pouvez-vous dire de la mesure $T_*^n(\mu(\cdot|[s_0, \ldots, s_{n-1}]))$?

(d) En déduire que, sous la loi μ, et pour une filtration bien choisie, le processus $(\sigma \circ T^n)$ est une chaîne de Markov.

(e) Peut-on réaliser toute chaîne de Markov sur un espace d'état fini de cette façon ?

4. **CODAGE SYMBOLIQUE D'UN AUTOMORPHISME DU TORE**

On considère dans cet exercice l'application du chat d'Arnol'd :

$$T : \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

qui agit sur le tore $\mathbb{T}^2$.

(a) Quel est le spectre de T ? En déduire que T préserve la mesure de Lebesgue et est mélangeante pour cette mesure.

(b) Quels sont les sous-espaces propres E^u et E^s de T (agissant sur $\mathbb{R}^2$) ? Construire deux feuilletages T-invariants sur $\mathbb{T}^2$.

On construit deux rectangles S_0 et S_1 qui recouvrent le tore de la façon suivante. Le rectangle S_0 (vu dans $\mathbb{R}^2$) a pour bords les droites E^u, $(1,1) + E^u$, $(0,1) + E^s$, E^s. Le rectangle S_1 (vu dans $\mathbb{R}^2$) a pour bords les droites E^u, $(1,0) + E^u$, E^s, $(1,1) + E^s$.

(c) Faire un dessin.

(d) Déterminer les composantes connexes de $T^{-1}(\mathring{S}_i) \cap \mathring{S}_j$. On les note $(R_k)_{k \in K}$.

(e) De même que dans l'exercice précédent, on note $R(x)$ l'élément de $(R_k)_{k \in K}$ auquel x appartient. Justifier le fait que $(R \circ T^n)_{n \geq 0}$ est une chaîne de Markov pour la mesure de Lebesgue sur $\mathbb{T}^2$ et une filtration bien choisie.

(f) Soit A la matrice $|K| \times |K|$ telle que :

$$A_{ij} = \begin{cases} 1 & \text{si} \quad R_i \cap T^{-1} R_j \neq \emptyset \\ 0 & \text{sinon} \end{cases}.$$

Écrire la matrice A. Quel est son rayon spectral ?