

TD 02 : Chaînes de Markov

# Propriétés de décorrélation des chaînes de Markov

## 1. CHAÎNES DE MARKOV

Soit  $\Sigma$  un ensemble fini de cardinalité au moins 2. Une *matrice stochastique* sur  $\Sigma$  est une matrice carrée, dont les lignes et les colonnes sont indicées par  $\Sigma$ , 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  $\Sigma$ , il existe  $n \geq 1$  tel que  $(A^n)_{ij} > 0$ , et *apériodique* si pour tous  $i, j$  dans  $\Sigma$ , il existe  $n_0 \geq 1$  tel que  $(A^n)_{ij} > 0$  pour tout  $n \geq n_0$ . Le théorème de Perron-Frobenius 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  $\rho$  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  $\rho$  est simple, et est une racine  $p$ -ième de l'unité.

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

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

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

- (a) Considérons une chaîne de Markov irréductible. Montrer qu'il existe une unique mesure de probabilité  $\mu$  sur  $\Sigma$  qui soit invariante au sens suivant : si  $M_0$  est de loi  $\mu$ , alors  $M_1$  l'est aussi.

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

- (b) Montrer que  $(\Omega, \bar{\mu}, T)$  est un système dynamique préservant la mesure.
- (c) Soient  $z \in \mathbb{S}_1$  et  $f \in \mathbb{L}^2(\Omega, \bar{\mu})$ . Montrer que  $\mathcal{L}(f) = zf$  si et seulement si  $f \circ T = \bar{z}f$ . *Indication : pour le sens direct, on pourra calculer  $\|f \circ T - \bar{z}f\|_{\mathbb{L}^2}$ .*
- (d) Soient  $f$  et  $g$  deux fonctions complexes sur  $\Sigma$ . Exprimer  $\mathbb{E}(f \cdot g \circ T^n)$  à l'aide de  $\mu$  et  $A^n$ . Si  $A$  est apériodique, que peut-on dire quand  $n$  tend vers l'infini ?
- (e) Pour tout  $N \geq 0$ , on note  $\mathcal{F}_N$  l'espace des fonctions complexes sur  $\Omega$  qui ne dépendent que des  $N$  premières coordonnées. 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  $g \in \mathbb{L}^1(\Omega, \bar{\mu})$ , 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.

- (f) En déduire que si  $A$  est apériodique, alors  $(\Omega, \bar{\mu}, T)$  est mélangeant.
- (g) On suppose maintenant que  $A$  est seulement irréductible, et on considère la valeur propre  $e^{i2\pi k/p}$ . Montrer qu'il existe un vecteur propre associé dont toutes les coordonnées sont des racines  $p$ -ièmes de l'unité. *Indication : utiliser la question c).*
- (h) En déduire qu'il existe une partition  $(\Sigma_k)_{k \in \mathbb{Z}/p\mathbb{Z}}$  de  $\Sigma$  telle que  $T$  envoie chaque  $[\Sigma_k]$  sur  $[\Sigma_{k+1}]$ . Que peut-on dire de  $A^p$  ?
- (i) En déduire que si  $A$  est irréductible, alors  $(\Omega, \bar{\mu}, T)$  est ergodique.
- (j) Que vaut l'entropie  $h_{\bar{\mu}}(T)$  ?

## 2. UN THÉORÈME CENTRAL LIMITE

Le but de cet exercice est de démontrer un théorème central limite pour des chaînes de Markov<sup>1</sup>. Soit  $\Sigma$  un ensemble fini, de cardinal au moins 2. Soit  $A$  la matrice de transition d'une chaîne de Markov apériodique sur  $A$ . Soit  $\mu$  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  $\Sigma$  telle que  $\mathbb{E}_\mu(F) = 0$ . Soit  $\mathcal{F}$  l'espace des fonctions complexes sur  $\Sigma$ . Pour tout  $\omega \in \mathbb{R}$ , on définit des opérateurs sur  $\mathcal{F}$  par :

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

<sup>1</sup>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 irrationnelles sur le cercle.

- (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  $\mathcal{F}$ ,

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

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

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

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

On admettra que, pour tout  $\omega$  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  $\mathbb{E}_{\bar{\mu}}(g(\omega)) = 1$ .

Soit  $P_{\omega}$  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  $\omega$  assez petit.

- (c) Que valent  $\lambda(0)$  et  $g(0)$  ?
- (d) Montrer que, pour tout  $\omega$  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  $\mathcal{F}_0 := \{f \in \mathcal{F} : \mathbb{E}_{\bar{\mu}}(f) = 0\}$ , 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  $\mathbb{E}_{\bar{\mu}}\left(e^{i\omega \frac{S_n^F}{\sqrt{n}}}\right)$  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  $\Sigma$  un ensemble fini. Soit  $(a_s)_{s \in \Sigma}$  une partition de  $(a, b)$  en sous-ensemble  $(a_s, b_s)$ . Soit  $T : (a, b) \rightarrow (a, b)$  une application. On suppose que<sup>2</sup> :

- $T$  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.
- (b) En déduire qu'il existe au moins une mesure de probabilité  $T$ -invariante  $\mu$ , 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  $(s_0, \dots, s_{n-1}) \in \Sigma^n$ . Supposons que  $[s_0, \dots, s_{n-1}] := \bigcap_{k=0}^{n-1} T^{-k}([a_{s_k}, b_{s_k}))$  est de  $\mu$ -mesure non nulle. Que pouvez-vous dire de la mesure  $T_*^n(\mu(\cdot|[s_0, \dots, s_{n-1}]))$  ?
- (d) En déduire que, sous la loi  $\mu$ , 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  $\mathbb{T}^2$ .

<sup>2</sup>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.

- (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  $T^{-1}(\dot{S}_i) \cap \dot{S}_j$ . 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 :

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

Écrire la matrice  $A$ . Quel est son rayon spectral ?