Matrices et Chaînes de Markov

Benoît Cadre, ENS Cachan Bretagne

Formation IPR 2012

27 sept. 2012

1. Le modèle d'Ehrenfest

- Echange gazeux entre 2 compartiments A et B

- Modélisation de la dynamique :
  - ▷ molécules numérotées $1, \cdots, N$
  - ▷ à chaque instant $n = 1, 2, \cdots$:
    1. tirage au hasard dans $\{1, \cdots, N\}$
    2. la molécule tirée est changée de compartiment

- Formalisation :
  - ▷ à l'instant n :
    - $X_n =$ v.a. qui représente le nombre de molécules en A
    - $U_{n+1} \sim \mathcal{U}\{1, \cdots, N\}$, indépendante de $X_0, \cdots, X_n$
  - ▷ description récursive :

$$X_{n+1} = \begin{cases} X_n + 1 & \text{si } U_{n+1} \text{ est dans } B \\ X_n - 1 & \text{si } U_{n+1} \text{ est dans } A \end{cases}$$

Valeur présente = F(valeur du passé immédiat , innovation)

ou bien

« **Sachant le passé, le présent ne dépend que du passé immédiat** »

- Formalisation :
  - ▷ à l'instant n :
    - X_n = v.a. qui représente le nombre de molécules en A
    - $U_{n+1} \sim \mathcal{U}\{1, \cdots, N\}$, indépendante de $X_0, \cdots, X_n$
  - ▷ description récursive :

$$X_{n+1} = \begin{cases} X_n + 1 & \text{si } U_{n+1} \text{ est dans } B \\ X_n - 1 & \text{si } U_{n+1} \text{ est dans } A \end{cases}$$

Valeur présente = F(valeur du passé immédiat , innovation)

ou bien

≪ Sachant le passé, le présent ne dépend que du passé immédiat ≫

- Formalisation :
  - ▷ à l'instant n :
    - $X_n =$ v.a. qui représente le nombre de molécules en A
    - $U_{n+1} \sim \mathcal{U}\{1, \cdots, N\}$, indépendante de $X_0, \cdots, X_n$
  - ▷ description récursive :

$$X_{n+1} = \begin{cases} X_n + 1 & \text{si } U_{n+1} \text{ est dans } B \\ X_n - 1 & \text{si } U_{n+1} \text{ est dans } A \end{cases}$$

Valeur présente = F(valeur du passé immédiat , innovation)

ou bien

« Sachant le passé, le présent ne dépend que du passé immédiat »

2. Hypothèse de Markov

- E = espace des états (**discret**) et $(X_n)_{n\geq 0}$, suite de v.a. à valeurs dans E

Définition

$(X_n)_{n\geq 0}$ est une chaîne de Markov (CDM) si pour tous n et $i_0, \cdots, i_n, j \in E$:

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \cdots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n)$$

$$\Longleftrightarrow$$

$$X_{n+1} = \textit{fonction}\ (X_n, U_{n+1}),\ \textit{avec}\ U_{n+1}\ \textit{indépendante de}\ X_0, \cdots, X_n.$$

2. Hypothèse de Markov

- E = espace des états (**discret**) et $(X_n)_{n\geq 0}$, suite de v.a. à valeurs dans E

Définition

$(X_n)_{n\geq 0}$ est une chaîne de Markov (CDM) si pour tous n et $i_0, \cdots, i_n, j \in E$:

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \cdots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n)$$

$$\Longleftrightarrow$$

$$X_{n+1} = \textit{fonction } (X_n, U_{n+1}), \ \textit{avec } U_{n+1} \textit{ indépendante de } X_0, \cdots, X_n.$$

- **Terminologie**
  - ▷ Lorsque les probabilités de transition $\mathbb{P}(X_{n+1} = j | X_n = i)$ sont indépendantes de n, la chaîne est dite homogène et

$$P = \big(P(i,j)\big)_{i,j \in E} = \big(\mathbb{P}(X_{n+1} = j | X_n = i)\big)_{i,j \in E} \text{ : matrice de transition}$$

  - ▷ P est une matrice stochastique :

$$\sum_{j \in E} P(i,j) = \sum_{j \in E} \mathbb{P}(X_{n+1} = j | X_n = i) = 1$$

- **Ex. chaîne d'Ehrenfest** : $(X_n)_{n \geq 0}$ CDMH d'espace d'états $E = \{0, \cdots, N\}$ et

$$P(i,j) = \mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} \frac{N-i}{N} & \text{si } j = i+1 \\ \frac{i}{N} & \text{si } j = i-1 \end{cases}$$

- **Terminologie**
  - ▷ Lorsque les probabilités de transition $\mathbb{P}(X_{n+1} = j | X_n = i)$ sont indépendantes de n, la chaîne est dite homogène et

$$P = \big(P(i,j)\big)_{i,j \in E} = \big(\mathbb{P}(X_{n+1} = j | X_n = i)\big)_{i,j \in E} \text{ : matrice de transition}$$

  - ▷ P est une matrice stochastique :

$$\sum_{j \in E} P(i,j) = \sum_{j \in E} \mathbb{P}(X_{n+1} = j | X_n = i) = 1$$

- **Ex. chaîne d'Ehrenfest** : $(X_n)_{n \geq 0}$ CDMH d'espace d'états $E = \{0, \cdots, N\}$ et

$$P(i,j) = \mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} \frac{N-i}{N} & \text{si } j = i+1 \\ \frac{i}{N} & \text{si } j = i-1 \end{cases}$$

- **Terminologie**
  - ▷ Lorsque les probabilités de transition $\mathbb{P}(X_{n+1} = j | X_n = i)$ sont indépendantes de n, la chaîne est dite homogène et

$$P = \big(P(i,j)\big)_{i,j \in E} = \big(\mathbb{P}(X_{n+1} = j | X_n = i)\big)_{i,j \in E} : \text{matrice de transition}$$

  - ▷ P est une matrice stochastique :

$$\sum_{j \in E} P(i,j) = \sum_{j \in E} \mathbb{P}(X_{n+1} = j | X_n = i) = 1$$

- **Ex. chaîne d'Ehrenfest** : $(X_n)_{n \geq 0}$ CDMH d'espace d'états $E = \{0, \cdots, N\}$ et

$$P(i,j) = \mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} \frac{N-i}{N} & \text{si } j = i+1 \\ \frac{i}{N} & \text{si } j = i-1 \end{cases}$$

• **Propriété** — $(X_n)_{n\geq 0}$ CDMH sur E, $P =$ matrice de transition (dorénavant).
Probabilité d'une trajectoire :

$$\mathbb{P}(X_0 = i_0, \cdots, X_n = i_n) = \mathbb{P}(X_0 = i_0) \prod_{k=0}^{n-1} P(i_k, i_{k+1})$$

▷ *loi du processus $(X_n)_{n\geq 0}$ caractérisée par P et la loi initiale (loi de X_0)*

• **Ex. chaîne d'Ehrenfest**, lorsque A est vide à l'instant initial ($X_0 = 0$).
Trajectoire la plus probable : $0, 1, 2, \cdots, N/2(\pm 1), N/2(\pm 1), \cdots$;
asymptotique (probable) de X_n de $N/2$ en moyenne :

$$\mathbb{E}X_{n+1} = \mathbb{E}\,\mathbb{E}[X_{n+1}|X_n] = 1 + \left(1 - \frac{2}{N}\right)\mathbb{E}X_n$$

• **Propriété** — $(X_n)_{n\geq 0}$ CDMH sur E, $P =$ matrice de transition (dorénavant).
Probabilité d'une trajectoire :

$$\mathbb{P}(X_0 = i_0, \cdots, X_n = i_n) = \mathbb{P}(X_0 = i_0) \prod_{k=0}^{n-1} P(i_k, i_{k+1})$$

▷ *loi du processus $(X_n)_{n\geq 0}$ caractérisée par P et la loi initiale (loi de X_0)*

• **Ex. chaîne d'Ehrenfest**, lorsque A est vide à l'instant initial ($X_0 = 0$).
Trajectoire la plus probable : $0, 1, 2, \cdots, N/2(\pm 1), N/2(\pm 1), \cdots$;
asymptotique (probable) de X_n de $N/2$ en moyenne :

$$\mathbb{E}X_{n+1} = \mathbb{E}\,\mathbb{E}[X_{n+1}|X_n] = 1 + \left(1 - \frac{2}{N}\right)\mathbb{E}X_n$$

- **Propriété** __ Relation de Chapman-Kolmogorov :

$$\mathbb{P}(X_n = j | X_0 = i) = P^n(i,j)$$

- **Ex.** : $\alpha, \beta \in]0,1[$, $E = \{a,b\}$ et

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Décomposition valeurs propres / vecteurs propres :

$$(\alpha+\beta)P^n = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{pmatrix} \begin{pmatrix} \beta & \alpha \\ 1 & -1 \end{pmatrix} \to \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}$$

▷ Phénomène d'oubli de la condition initiale. Par exemple,

$$\lim_{n\to\infty} \mathbb{P}(X_n = a | X_0 = i) = \lim_{n\to\infty} \mathbb{P}(X_n = a) = \beta/(\alpha+\beta)$$

- **Propriété** — Relation de Chapman-Kolmogorov :

$$\mathbb{P}(X_n = j | X_0 = i) = P^n(i,j)$$

- **Ex.** : $\alpha, \beta \in]0,1[$, $E = \{a, b\}$ et

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Décomposition valeurs propres / vecteurs propres :

$$(\alpha+\beta)P^n = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{pmatrix} \begin{pmatrix} \beta & \alpha \\ 1 & -1 \end{pmatrix} \to \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}$$

▷ Phénomène d'oubli de la condition initiale. Par exemple,

$$\lim_{n\to\infty} \mathbb{P}(X_n = a | X_0 = i) = \lim_{n\to\infty} \mathbb{P}(X_n = a) = \beta/(\alpha+\beta)$$

- **Propriété** __ Relation de Chapman-Kolmogorov :

$$\mathbb{P}(X_n = j | X_0 = i) = P^n(i,j)$$

- **Ex.** : $\alpha, \beta \in]0,1[$, $E = \{a, b\}$ et

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Décomposition valeurs propres / vecteurs propres :

$$(\alpha+\beta)P^n = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{pmatrix} \begin{pmatrix} \beta & \alpha \\ 1 & -1 \end{pmatrix} \to \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}$$

▷ Phénomène d'oubli de la condition initiale. Par exemple,

$$\lim_{n\to\infty} \mathbb{P}(X_n = a | X_0 = i) = \lim_{n\to\infty} \mathbb{P}(X_n = a) = \beta/(\alpha+\beta)$$

3. Communication entre états

- Propagation de la CDMH $(X_n)_{n\geq 0}$? si tous les états sont visités, d'où :

Définition

Les états i et j communiquent (noté $i \leftrightarrow j$) si il existe $n, k \geq 0$ tels que

$$P^n(i,j) = \mathbb{P}(X_n = j | X_0 = i) > 0 \text{ et } P^k(j,i) = \mathbb{P}(X_k = i | X_0 = j) > 0$$

- **Exemple** : tous les états de la chaîne ci-dessous ne communiquent pas

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

$E = \{a, b, c\}$ se décompose en 2 ensembles $\{a\}$ et $\{b, c\}$ aux statuts différents

3. Communication entre états

• Propagation de la CDMH $(X_n)_{n\geq 0}$? si tous les états sont visités, d'où :

Définition

Les états i et j communiquent (noté $i \leftrightarrow j$) si il existe $n, k \geq 0$ tels que

$$P^n(i,j) = \mathbb{P}(X_n = j | X_0 = i) > 0 \text{ et } P^k(j,i) = \mathbb{P}(X_k = i | X_0 = j) > 0$$

• **Exemple** : tous les états de la chaîne ci-dessous ne communiquent pas

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

$E = \{a, b, c\}$ se décompose en 2 ensembles $\{a\}$ et $\{b, c\}$ aux statuts différents

3. Communication entre états

- Propagation de la CDMH $(X_n)_{n\geq 0}$? si tous les états sont visités, d'où :

Définition

Les états i et j communiquent (noté $i \leftrightarrow j$) si il existe $n, k \geq 0$ tels que

$$P^n(i,j) = \mathbb{P}(X_n = j | X_0 = i) > 0 \text{ et } P^k(j,i) = \mathbb{P}(X_k = i | X_0 = j) > 0$$

- **Exemple** : tous les états de la chaîne ci-dessous ne communiquent pas

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

$E = \{a, b, c\}$ se décompose en 2 ensembles $\{a\}$ et $\{b, c\}$ aux statuts différents

- $\leftrightarrow$ est une relation d'équivalence et

$$E = \bigsqcup_{\ell \geq 1} C_\ell, \;\; C_\ell \text{ classe d'équivalence d'états}$$

  ▷ la chaîne est dite irréductible lorsqu'il n'y a qu'une classe d'états

- **Ex.** : la chaîne d'Ehrenfest est irréductible

- $\leftrightarrow$ est une relation d'équivalence et

$$E = \bigsqcup_{\ell \geq 1} C_\ell, \;\; C_\ell \text{ classe d'équivalence d'états}$$

  ▷ la chaîne est dite irréductible lorsqu'il n'y a qu'une classe d'états

- **Ex.** : la chaîne d'Ehrenfest est irréductible

4. Classification des états

- Dynamique dans chaque classe d'équivalence ?
- Le temps d'atteinte de $(X_n)_{n\geq 0}$ dans l'état $i \in E$ est :

$$T_i = \inf\{n \geq 1 \ : \ X_n = i\}, \text{ avec } \inf \emptyset = \infty$$

Définition

i est récurrent (resp. transitoire) si $\mathbb{P}(T_i < \infty | X_0 = i) = 1$ (resp. < 1)

- **Ex.** : Pour la chaîne d'espace d'états $E = \{a, b, c\}$ et

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \Rightarrow a \text{ transitoire, } b, c \text{ récurrents}$$

4. Classification des états

- Dynamique dans chaque classe d'équivalence ?
- Le temps d'atteinte de $(X_n)_{n\geq 0}$ dans l'état $i \in E$ est :

$$T_i = \inf\{n \geq 1 \; : \; X_n = i\}, \text{ avec } \inf \emptyset = \infty$$

Définition

i est récurrent (resp. transitoire) si $\mathbb{P}(T_i < \infty | X_0 = i) = 1$ *(resp.* < 1*)*

- **Ex.** : Pour la chaîne d'espace d'états $E = \{a, b, c\}$ et

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \Rightarrow a \text{ transitoire, } b, c \text{ récurrents}$$

4. Classification des états

- Dynamique dans chaque classe d'équivalence ?
- Le temps d'atteinte de $(X_n)_{n\geq 0}$ dans l'état $i \in E$ est :

$$T_i = \inf\{n \geq 1 \,:\, X_n = i\}, \text{ avec } \inf \emptyset = \infty$$

Définition

i est récurrent (resp. transitoire) si $\mathbb{P}(T_i < \infty | X_0 = i) = 1$ *(resp.* < 1*)*

- **Ex.** : Pour la chaîne d'espace d'états $E = \{a, b, c\}$ et

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \Rightarrow a \text{ transitoire, } b, c \text{ récurrents}$$

- Deux phénomènes apparaissent avec cet exemple :
  - ▷ dans une même classe, les états sont de même nature

$$i \leftrightarrow j \Rightarrow i, j \text{ sont de même nature}$$

Cas particulier : $|E| < \infty$ + chaîne irréductible $\Rightarrow$ états récurrents

  - ▷ la chaîne visite une infinité de fois un état récurrent (prop. Markov fort)

$$i \text{ récurrent} \Leftrightarrow \mathbb{P}\Big(\sum_n \mathbf{1}_{\{X_n = i\}} = \infty \Big| X_0 = i \Big) = 1$$

- Deux phénomènes apparaissent avec cet exemple :
  - ▷ dans une même classe, les états sont de même nature

$$i \leftrightarrow j \Rightarrow i, j \text{ sont de même nature}$$

Cas particulier : $|E| < \infty$ + chaîne irréductible $\Rightarrow$ états récurrents

  - ▷ la chaîne visite une infinité de fois un état récurrent (prop. Markov fort)

$$i \text{ récurrent} \Leftrightarrow \mathbb{P}\Big(\sum_n \mathbf{1}_{\{X_n = i\}} = \infty \Big| X_0 = i \Big) = 1$$

- Récurrence/matrice de transition : ≪ i récurrent si $P^n(i,i)$ grand ≫

$$i \text{ est récurrent } \Leftrightarrow \sum_n P^n(i,i) = \infty$$

- **Ex. : marche aléatoire simple sur $E = \mathbb{Z}^d$**

$$\text{Matrice de transition : } P(i,j) = \begin{cases} 1/(2d) & \text{si } |i-j| = 1 \\ 0 & \text{sinon} \end{cases}$$

$$P^{2k+1}(0,0) = 0 \text{ et } \begin{cases} P^{2k}(0,0) \sim C/k^{d/2} & \text{si } d \leq 2 \\ P^{2k}(0,0) \leq C/k^{d/2} & \text{si } d \geq 3 \end{cases}$$

Par irréductibilité, tous les états sont récurrents si, et seulement si $d \leq 2$

- Récurrence/matrice de transition : ≪ i récurrent si $P^n(i,i)$ grand ≫

$$i \text{ est récurrent } \Leftrightarrow \sum_n P^n(i,i) = \infty$$

- **Ex. : marche aléatoire simple sur $E = \mathbb{Z}^d$**

$$\text{Matrice de transition : } P(i,j) = \begin{cases} 1/(2d) & \text{si } |i-j| = 1 \\ 0 & \text{sinon} \end{cases}$$

$$P^{2k+1}(0,0) = 0 \text{ et } \begin{cases} P^{2k}(0,0) \sim C/k^{d/2} & \text{si } d \leq 2 \\ P^{2k}(0,0) \leq C/k^{d/2} & \text{si } d \geq 3 \end{cases}$$

Par irréductibilité, tous les états sont récurrents si, et seulement si $d \leq 2$

- Récurrence/matrice de transition : $\ll i$ récurrent si $P^n(i,i)$ grand $\gg$

  i est récurrent $\Leftrightarrow \sum_n P^n(i,i) = \infty$

- **Ex. : marche aléatoire simple sur $E = \mathbb{Z}^d$**

$$\text{Matrice de transition : } P(i,j) = \begin{cases} 1/(2d) & \text{si } |i-j| = 1 \\ 0 & \text{sinon} \end{cases}$$

$$P^{2k+1}(0,0) = 0 \text{ et } \begin{cases} P^{2k}(0,0) \sim C/k^{d/2} & \text{si } d \leq 2 \\ P^{2k}(0,0) \leq C/k^{d/2} & \text{si } d \geq 3 \end{cases}$$

Par irréductibilité, tous les états sont récurrents si, et seulement si $d \leq 2$

5. Loi stationnaire

- **Battage de cartes** __ Jeu de N cartes
  - ▷ Comment caractériser la performance d'une méthode de battage ?
  - ▷ Modèle :
    - → espace d'états E = permutations de $\{1, \cdots, N\}$
    - → X_n = état du paquet à la n-ème étape du battage
    - → restriction : $(X_n)_{n \geq 0}$ CDMH sur E (ex. : battage par insertion,...)

5. Loi stationnaire

- **Battage de cartes** __ Jeu de N cartes
  - ▷ Comment caractériser la performance d'une méthode de battage ?
  - ▷ Modèle :
    - $\rightarrow$ espace d'états $E =$ permutations de $\{1, \cdots, N\}$
    - $\rightarrow$ $X_n =$ état du paquet à la n-ème étape du battage
    - $\rightarrow$ restriction : $(X_n)_{n\geq 0}$ CDMH sur E (ex. : battage par insertion,...)

▷ Méthode performante si loi$(X_n) \approx \mathcal{U}(E)$, i.e.

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \frac{1}{N!}$$

Or, on a aussi (formule des probas totales) :

$$\mathbb{P}(X_{n+1} = j) = \sum_{i\in E} P(i,j)\mathbb{P}(X_n = i)$$

▷ Une méthode de battage (markovienne) doit vérifier

$$\pi = \frac{1}{N!}(1\cdots 1) \text{ seule probabilité solution de } \pi = \pi P$$

▷ Méthode performante si loi(X_n) $\approx \mathcal{U}(E)$, i.e.

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \frac{1}{N!}$$

Or, on a aussi (formule des probas totales) :

$$\mathbb{P}(X_{n+1} = j) = \sum_{i\in E} P(i,j)\mathbb{P}(X_n = i)$$

▷ Une méthode de battage (markovienne) doit vérifier

$$\pi = \frac{1}{N!}(1\cdots 1) \text{ seule probabilité solution de } \pi = \pi P$$

- **Convention** : probabilité π sur E sous la forme d'un vecteur ligne,

$$\pi = (\pi(i), i \in E) \text{ avec } \pi(i) \geq 0 \text{ et } \sum_{i \in E} \pi(i) = 1$$

- La loi de X_n sous la loi initiale $\pi = \text{loi}(X_0)$ est πP^n car

$$\mathbb{P}(X_n = i) = \sum_{\ell \in E} \pi(\ell) P^n(\ell, i)$$

▷ si π probabilité telle que $\pi P = \pi$, $\text{loi}(X_0) = \pi \Rightarrow \text{loi}(X_n) = \pi P^n = \pi$

Définition

Une loi stationnaire est une probabilité π sur E telle que $\pi = \pi P$

▷ existe dès que $|E| < \infty$, unique si la chaîne est irréductible

- **Convention** : probabilité π sur E sous la forme d'un vecteur ligne,

$$\pi = (\pi(i), i \in E) \text{ avec } \pi(i) \geq 0 \text{ et } \sum_{i\in E} \pi(i) = 1$$

- La loi de X_n sous la loi initiale $\pi = \text{loi}(X_0)$ est πP^n car

$$\mathbb{P}(X_n = i) = \sum_{\ell \in E} \pi(\ell) P^n(\ell, i)$$

▷ si π probabilité telle que $\pi P = \pi$, $\text{loi}(X_0) = \pi \Rightarrow \text{loi}(X_n) = \pi P^n = \pi$

Définition

Une loi stationnaire est une probabilité π sur E telle que $\pi = \pi P$

▷ existe dès que $|E| < \infty$, unique si la chaîne est irréductible

- **Convention** : probabilité π sur E sous la forme d'un vecteur ligne,

$$\pi = (\pi(i), i \in E) \text{ avec } \pi(i) \geq 0 \text{ et } \sum_{i \in E} \pi(i) = 1$$

- La loi de X_n sous la loi initiale $\pi = \text{loi}(X_0)$ est πP^n car

$$\mathbb{P}(X_n = i) = \sum_{\ell \in E} \pi(\ell) P^n(\ell, i)$$

▷ si π probabilité telle que $\pi P = \pi$, $\text{loi}(X_0) = \pi \Rightarrow \text{loi}(X_n) = \pi P^n = \pi$

Définition

Une loi stationnaire est une probabilité π sur E telle que $\pi = \pi P$

▷ existe dès que $|E| < \infty$, unique si la chaîne est irréductible

- **Convention** : probabilité π sur E sous la forme d'un vecteur ligne,

$$\pi = (\pi(i), i \in E) \text{ avec } \pi(i) \geq 0 \text{ et } \textstyle\sum_{i \in E} \pi(i) = 1$$

- La loi de X_n sous la loi initiale $\pi = \text{loi}(X_0)$ est πP^n car

$$\mathbb{P}(X_n = i) = \sum_{\ell \in E} \pi(\ell) P^n(\ell, i)$$

▷ si π probabilité telle que $\pi P = \pi$, $\text{loi}(X_0) = \pi \Rightarrow \text{loi}(X_n) = \pi P^n = \pi$

Définition

Une loi stationnaire est une probabilité π sur E telle que $\pi = \pi P$

▷ existe dès que $|E| < \infty$, unique si la chaîne est irréductible

- **Contre-ex.** :

  ▷ $E = \mathbb{N}$ et $P(i, i+1) = 1$: pas de loi stationnaire

  ▷ $E = \{a, b\}$ et $P = \mathrm{Id}$: toute probabilité est loi stationnaire

- **Ex. chaîne d'Ehrenfest** : si $\pi = \mathit{Binomiale}(N, 1/2)$, alors

$$\pi(j)P(j,i) = \pi(i)P(i,j) \text{ (équation de réversibilité)}$$

donc π est la seule loi stationnaire de la chaîne d'Ehrenfest. Si loi$(X_0) = \pi$,

$$\mathbb{E}X_n = N/2 \text{ et } \mathrm{var}(X_n) = N/4$$

- **Contre-ex.** :

  $\triangleright$ $E = \mathbb{N}$ et $P(i, i+1) = 1$: pas de loi stationnaire

  $\triangleright$ $E = \{a, b\}$ et $P = \mathrm{Id}$: toute probabilité est loi stationnaire

- **Ex. chaîne d'Ehrenfest** : si $\pi = \mathit{Binomiale}(N, 1/2)$, alors

$$\pi(j)P(j,i) = \pi(i)P(i,j) \text{ (équation de réversibilité)}$$

donc π est la seule loi stationnaire de la chaîne d'Ehrenfest. Si loi$(X_0) = \pi$,

$$\mathbb{E}X_n = N/2 \text{ et } \mathrm{var}(X_n) = N/4$$

6. Périodicité

- **Dorénavant, $|E| < \infty$**
- **Problème** : asymptotique de $(X_n)_{n \geq 0}$
  - ▷ liée à celle de P^n (relation de Chapman-Kolmogorov)
  - ▷ $(P^n)_n$ compact $\Rightarrow$ il existe une sous-suite convergente
  - ▷ description de la sous-suite ?

6. Périodicité

- **Dorénavant,** $|E| < \infty$
- **Problème** : asymptotique de $(X_n)_{n\geq 0}$
  - ▷ liée à celle de P^n (relation de Chapman-Kolmogorov)
  - ▷ $(P^n)_n$ compact $\Rightarrow$ il existe une sous-suite convergente
  - ▷ description de la sous-suite ?

6. Périodicité

- **Dorénavant, $|E| < \infty$**
- **Problème** : asymptotique de $(X_n)_{n\geq 0}$
  - ▷ liée à celle de P^n (relation de Chapman-Kolmogorov)
  - ▷ $(P^n)_n$ compact $\Rightarrow$ il existe une sous-suite convergente
  - ▷ description de la sous-suite ?

6. Périodicité

- **Dorénavant, $|E| < \infty$**
- **Problème** : asymptotique de $(X_n)_{n\geq 0}$
  - ▷ liée à celle de P^n (relation de Chapman-Kolmogorov)
  - ▷ $(P^n)_n$ compact $\Rightarrow$ il existe une sous-suite convergente
  - ▷ description de la sous-suite ?

- **Ex. chaîne d'Ehrenfest** :

$$P^{2n+1}(0,0) = 0 \text{ et } \liminf_{n\to\infty} P^{2n}(0,0) > 0$$

- ▷ sous la condition initiale $X_0 = 0$, $(X_n)_{n\geq 0}$ ne converge pas en loi, mais on peut montrer que la sous-suite $(X_{2n})_{n\geq 0}$ converge
- ▷ or, $(X_{2n})_{n\geq 0}$ CDM homogène de matrice de transition P^2
- ▷ *différence entre ces 2 CDM* : dynamique des retours sur un état, e. g. 0

  → Cas de $(X_n)_{n\geq 0}$: retour en 0 que pour des instants multiples de 2

  → Cas de $(X_{2n})_{n\geq 0}$: mélange

- **Ex. chaîne d'Ehrenfest** :

$$P^{2n+1}(0,0) = 0 \text{ et } \liminf_{n\to\infty} P^{2n}(0,0) > 0$$

  - ▷ sous la condition initiale $X_0 = 0$, $(X_n)_{n\geq 0}$ ne converge pas en loi, mais on peut montrer que la sous-suite $(X_{2n})_{n\geq 0}$ converge
  - ▷ or, $(X_{2n})_{n\geq 0}$ CDM homogène de matrice de transition P^2
  - ▷ *différence entre ces 2 CDM* : dynamique des retours sur un état, e. g. 0

    → Cas de $(X_n)_{n\geq 0}$: retour en 0 que pour des instants multiples de 2

    → Cas de $(X_{2n})_{n\geq 0}$: mélange

- **Ex. chaîne d'Ehrenfest** :

$$P^{2n+1}(0,0) = 0 \text{ et } \liminf_{n\to\infty} P^{2n}(0,0) > 0$$

  - ▷ sous la condition initiale $X_0 = 0$, $(X_n)_{n\geq 0}$ ne converge pas en loi, mais on peut montrer que la sous-suite $(X_{2n})_{n\geq 0}$ converge
  - ▷ or, $(X_{2n})_{n\geq 0}$ CDM homogène de matrice de transition P^2
  - ▷ *différence entre ces 2 CDM* : dynamique des retours sur un état, e. g. 0

    $\rightarrow$ Cas de $(X_n)_{n\geq 0}$: retour en 0 que pour des instants multiples de 2

    $\rightarrow$ Cas de $(X_{2n})_{n\geq 0}$: mélange

Définition

La période de l'état i est

$$d(i) = \mathrm{pgcd}\{n \geq 1 \ : \ P^n(i,i) > 0\}$$

i est apériodique si $d(i) = 1$

- **Ex.** : $E = \{a, b, c, d\}$ et

$$P = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad : \quad d(i) = 3$$

- **Ex. chaîne d'Ehrenfest** : $d(i) = 2$ pour $(X_n)_{n\geq 0}$, $= 1$ pour $(X_{2n})_{n\geq 0}$

Définition

La période de l'état i est

$$d(i) = \mathrm{pgcd}\{n \geq 1 \ : \ P^n(i,i) > 0\}$$

i est apériodique si $d(i) = 1$

- **Ex.** : $E = \{a, b, c, d\}$ et

$$P = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad : \quad d(i) = 3$$

- **Ex. chaîne d'Ehrenfest** : $d(i) = 2$ pour $(X_n)_{n\geq 0}$, $= 1$ pour $(X_{2n})_{n\geq 0}$

Définition

*La **période** de l'état i est*

$$d(i) = \mathrm{pgcd}\{n \geq 1 \;:\; P^n(i,i) > 0\}$$

*i est **apériodique** si $d(i) = 1$*

- **Ex.** : $E = \{a, b, c, d\}$ et

$$P = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad : \quad d(i) = 3$$

- **Ex. chaîne d'Ehrenfest** : $d(i) = 2$ pour $(X_n)_{n\geq 0}$, $= 1$ pour $(X_{2n})_{n\geq 0}$

- **Propriété** __ la période est une caractéristique de classe :

$$i \leftrightarrow j \Rightarrow d(i) = d(j)$$

▷ Décomposition de l'espace des états :

$$E = T \bigsqcup \bigsqcup_{\ell=1}^{k} C_\ell$$

avec T états transitoires, C_ℓ états récurrents de même période

- **Propriété** __ la période est une caractéristique de classe :

$$i \leftrightarrow j \Rightarrow d(i) = d(j)$$

▷ Décomposition de l'espace des états :

$$E = T \bigsqcup \bigsqcup_{\ell=1}^{k} C_\ell$$

avec T états transitoires, C_ℓ états récurrents de même période

• **Propriété** __ si la chaîne est irréductible et apériodique, alors P est primitive i.e.

$$P^n > 0 \text{ pour tout } n \text{ assez grand}$$

• **Contre-ex. chaîne d'Ehrenfest**, pour $N = 4$:

$$P^{2n+1} = \begin{pmatrix} 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \end{pmatrix} \text{ et } P^{2n} = \begin{pmatrix} \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \end{pmatrix}$$

• **Propriété** __ si la chaîne est irréductible et apériodique, alors P est primitive i.e.

$$P^n > 0 \text{ pour tout } n \text{ assez grand}$$

• **Contre-ex. chaîne d'Ehrenfest**, pour $N = 4$:

$$P^{2n+1} = \begin{pmatrix} 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \end{pmatrix} \text{ et } P^{2n} = \begin{pmatrix} \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \\ 0 & \star & 0 & \star & 0 \\ \star & 0 & \star & 0 & \star \end{pmatrix}$$

7. Asymptotique

Théorème de Perron-Frobenius : *le rayon spectral d'une matrice primitive est valeur propre simple dominante*

Comme le rayon spectral d'une matrice stochastique vaut 1,

1 est valeur propre simple dominante de la matrice de transition d'une chaîne irréductible et apériodique

7. Asymptotique

Théorème de Perron-Frobenius : *le rayon spectral d'une matrice primitive est valeur propre simple dominante*

Comme le rayon spectral d'une matrice stochastique vaut 1,

1 est valeur propre simple dominante de la matrice de transition d'une chaîne irréductible et apériodique

7. Asymptotique

Théorème de Perron-Frobenius : *le rayon spectral d'une matrice primitive est valeur propre simple dominante*

Comme le rayon spectral d'une matrice stochastique vaut 1,

1 est valeur propre simple dominante de la matrice de transition d'une chaîne irréductible et apériodique

Théorème

Supposons que la CDMH est irréductible, apériodique, de loi stationnaire π. Soit $\rho < 1$ le 2nd plus grand module des valeurs propres de P, et m la multiplicité. Alors,

$$\max_{j \in E} \left| \mathbb{P}(X_n = j | X_0 = i) - \pi(j) \right| = \max_{j \in E} \left| P^n(i,j) - \pi(j) \right| = O\left(n^{m-1} \rho^n\right)$$

- **Remarques**
  - ▷ oubli de la condition initiale
  - ▷ proche du régime stationnaire après un ≪ temps de chauffe ≫ très court

Théorème

Supposons que la CDMH est irréductible, apériodique, de loi stationnaire π. Soit $\rho < 1$ le 2nd plus grand module des valeurs propres de P, et m la multiplicité. Alors,

$$\max_{j \in E} \left| \mathbb{P}(X_n = j | X_0 = i) - \pi(j) \right| = \max_{j \in E} \left| P^n(i,j) - \pi(j) \right| = O\left(n^{m-1} \rho^n\right)$$

- **Remarques**
  - ▷ oubli de la condition initiale
  - ▷ proche du régime stationnaire après un « temps de chauffe » très court

Preuve __ cas P de valeurs propres distinctes

Etape 1 : P^n converge à vitesse exponentielle

$\triangleright$ $|E| = \ell + 1$, $1 > |\lambda_1| \geq \cdots \geq |\lambda_\ell|$

$$\lim_{n\to\infty} P^n = \lim_{n\to\infty} A^{-1} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \lambda_1^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_\ell^n \end{pmatrix} A = A^{-1} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} A$$

à vitesse exponentielle

Etape 2 : identification de la limite

▷ π^j proba. sur E et

$$\prod = \begin{pmatrix} \pi^1 \\ \vdots \\ \pi^{\ell+1} \end{pmatrix} = \lim_n P^n$$

▷ $\prod = \prod P$ (car $\mathrm{P}^{n+1} = P^n P$) $\Rightarrow \pi^j = \pi^j P$

$\Rightarrow \pi^j = \pi$ par unicité de la loi stationnaire

8. Voyageur de commerce

- **Problème** (simplifié)
  - ▷ Déterminer le chemin le plus court de A vers B
  - ▷ Pas d'algorithme déterministe en complexité polynomiale
- **Approche probabiliste**
  - ▷ $E =$ ensemble des trajectoires i de A vers B
  - ▷ $H(i) = \mathrm{longueur}(i) -$ longueur minimale des chemins $A \to B$
  - ▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\Big(- \frac{H(i)}{T} \Big) = \frac{1}{Z'_T} \exp\Big(- \frac{\text{longueur}(i)}{T} \Big)$$

avec $Z_T, Z'_T =$ constantes de normalisation, $T =$ température

8. Voyageur de commerce

- **Problème** (simplifié)
  - ▷ Déterminer le chemin le plus court de A vers B
  - ▷ Pas d'algorithme déterministe en complexité polynomiale
- **Approche probabiliste**
  - ▷ $E =$ ensemble des trajectoires i de A vers B
  - ▷ $H(i) = \mathrm{longueur}(i) -$ longueur minimale des chemins $A \to B$
  - ▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\left(-\frac{H(i)}{T}\right) = \frac{1}{Z'_T} \exp\left(-\frac{\mathrm{longueur}(i)}{T}\right)$$

avec $Z_T, Z'_T =$ constantes de normalisation, $T =$ température

8. Voyageur de commerce

- **Problème** (simplifié)
  - ▷ Déterminer le chemin le plus court de A vers B
  - ▷ Pas d'algorithme déterministe en complexité polynomiale
- **Approche probabiliste**
  - ▷ $E =$ ensemble des trajectoires i de A vers B
  - ▷ $H(i) = \mathrm{longueur}(i) -$ longueur minimale des chemins $A \to B$
  - ▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\Big(-\frac{H(i)}{T}\Big) = \frac{1}{Z'_T} \exp\Big(-\frac{\text{longueur}(i)}{T}\Big)$$

avec $Z_T, Z'_T =$ constantes de normalisation, $T =$ température

▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\Big(- \frac{H(i)}{T}\Big) = \frac{1}{Z'_T} \exp\Big(- \frac{\text{longueur}(i)}{T}\Big)$$

▷ *une observation* : si $T \approx 0$, une réalisation de π_T est un chemin (presque) optimal, car

$$\pi_T \approx \mathcal{U}\big(H^{-1}(0)\big)$$

▷ *idée à exploiter* : générer une réalisation de la loi π_T, pour $T \approx 0$

$\rightarrow$ ... mais Z_T, Z'_T impossibles à calculer en général

$\rightarrow$ algorithme de Métropolis-Hasting

▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\Big(- \frac{H(i)}{T}\Big) = \frac{1}{Z'_T} \exp\Big(- \frac{\text{longueur}(i)}{T}\Big)$$

▷ *une observation* : si $T \approx 0$, une réalisation de π_T est un chemin (presque) optimal, car

$$\pi_T \approx \mathcal{U}\big(H^{-1}(0)\big)$$

▷ *idée à exploiter* : générer une réalisation de la loi π_T, pour $T \approx 0$

$\rightarrow$ ... mais Z_T, Z'_T impossibles à calculer en général

$\rightarrow$ algorithme de Métropolis-Hasting

▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\Big(- \frac{H(i)}{T} \Big) = \frac{1}{Z'_T} \exp\Big(- \frac{\text{longueur}(i)}{T} \Big)$$

▷ *une observation* : si $T \approx 0$, une réalisation de π_T est un chemin (presque) optimal, car

$$\pi_T \approx \mathcal{U}\big(H^{-1}(0)\big)$$

▷ *idée à exploiter* : générer une réalisation de la loi π_T, pour $T \approx 0$

$\rightarrow$ ... mais Z_T, Z'_T impossibles à calculer en général

$\rightarrow$ algorithme de Métropolis-Hasting

▷ Probabilité de Gibbs :

$$\pi_T(i) = \frac{1}{Z_T} \exp\Big(- \frac{H(i)}{T} \Big) = \frac{1}{Z'_T} \exp\Big(- \frac{\text{longueur}(i)}{T} \Big)$$

▷ *une observation* : si $T \approx 0$, une réalisation de π_T est un chemin (presque) optimal, car

$$\pi_T \approx \mathcal{U}\big(H^{-1}(0)\big)$$

▷ *idée à exploiter* : générer une réalisation de la loi π_T, pour $T \approx 0$

$\rightarrow$ ... mais Z_T, Z'_T impossibles à calculer en général

$\rightarrow$ algorithme de Métropolis-Hasting

- **Outils de l'algorithme de Métropolis-Hasting** :
  - ▷ Matrice de sélection Q sur E = matrice stochastique telle que $Q(i,j) > 0$
  - ▷ Matrice d'acceptation ρ, définie par

$$\rho(i,j) = \min\left(1, \frac{\pi_T(j)Q(j,i)}{\pi_T(i)Q(i,j)}\right) \ : \ \text{ne fait pas intervenir } Z'_T\,!$$

- **Algorithme MH** :
  - ▷ $i_0 \in E$
  - ▷ *Etape* n : j_{n+1} = réalisation de $Q(i_n,.)$, u = réalisation de $\mathcal{U}[0,1]$

$$i_{n+1} = \begin{cases} j_{n+1} & \text{si } u \leq \rho(i_n, j_{n+1}) : \textit{transition acceptée} \\ i_n & \text{si } u > \rho(i_n, j_{n+1}) : \textit{transition refusée} \end{cases}$$

- **Outils de l'algorithme de Métropolis-Hasting** :
  - ▷ Matrice de sélection Q sur E = matrice stochastique telle que $Q(i,j) > 0$
  - ▷ Matrice d'acceptation ρ, définie par

$$\rho(i,j) = \min\left(1, \frac{\pi_T(j)Q(j,i)}{\pi_T(i)Q(i,j)}\right) \ : \ \text{ne fait pas intervenir } Z'_T\,!$$

- **Algorithme MH** :
  - ▷ $i_0 \in E$
  - ▷ *Etape n* : j_{n+1} = réalisation de $Q(i_n, .)$, u = réalisation de $\mathcal{U}[0,1]$

$$i_{n+1} = \begin{cases} j_{n+1} & \text{si } u \leq \rho(i_n, j_{n+1}) : \textit{transition acceptée} \\ i_n & \text{si } u > \rho(i_n, j_{n+1}) : \textit{transition refusée} \end{cases}$$

- $(i_n)_{n\geq 0}$ réalisation d'une CDMH $(X_n)_{n\geq 0}$ sur E de matrice de transition

$$P(i,j) = Q(i,j)\rho(i,j) \text{ si } i \neq j$$

- **Propriétés de** $(X_n)_{n\geq 0}$:
  - ▷ irréductible car $Q(i,j) > 0$
  - ▷ $|E| < \infty$, donc récurrente et admet une unique loi stationnaire
  - ▷ réversible par rapport à π_T :

$$P(i,j)\pi_T(i) = P(j,i)\pi_T(j)$$

  donc π_T est sa loi stationnaire
  - ▷ apériodique, car $P^2(i,i) > 0$ et $P^3(i,i) > 0$

- $(i_n)_{n\geq 0}$ réalisation d'une CDMH $(X_n)_{n\geq 0}$ sur E de matrice de transition

$$P(i,j) = Q(i,j)\rho(i,j) \text{ si } i \neq j$$

- **Propriétés de** $(X_n)_{n\geq 0}$:
  - ▷ irréductible car $Q(i,j) > 0$
  - ▷ $|E| < \infty$, donc récurrente et admet une unique loi stationnaire
  - ▷ réversible par rapport à π_T :

$$P(i,j)\pi_T(i) = P(j,i)\pi_T(j)$$

  donc π_T est sa loi stationnaire

  - ▷ apériodique, car $P^2(i,i) > 0$ et $P^3(i,i) > 0$

- $(i_n)_{n\geq 0}$ réalisation d'une CDMH $(X_n)_{n\geq 0}$ sur E de matrice de transition

$$P(i,j) = Q(i,j)\rho(i,j) \text{ si } i \neq j$$

- **Propriétés de** $(X_n)_{n\geq 0}$:
  - ▷ irréductible car $Q(i,j) > 0$
  - ▷ $|E| < \infty$, donc récurrente et admet une unique loi stationnaire
  - ▷ réversible par rapport à π_T :

$$P(i,j)\pi_T(i) = P(j,i)\pi_T(j)$$

  donc π_T est sa loi stationnaire

  - ▷ apériodique, car $P^2(i,i) > 0$ et $P^3(i,i) > 0$

- **Conclusion** :

  ▷ $\text{loi}(X_n) \approx \pi_T$ (vitesse exponentielle)

  ▷ $\pi_T \approx \mathcal{U}\big(\text{trajectoires optimales de A vers B}\big)$ (si $T \approx 0$)

  ▷ $i_n \approx$ **chemin optimal**