

Introduction aux chaînes de Markov (1)

Dynamique des lois, dynamique des observables, mesures stationnaires

Jérémie Unterberger

Nancy, septembre 2019

Biblio.:

- Olle Häggström, *Finite Markov chains and algorithmic applications*
- Carl Graham, *Chaînes de Markov. Cours et exercices corrigés*, Dunod
- James Norris, *Markov chains*

Biblio.:

- Olle Häggström, *Finite Markov chains and algorithmic applications*
- Carl Graham, *Chaînes de Markov. Cours et exercices corrigés*, Dunod
- James Norris, *Markov chains*
- Bernard Ycart, *Modèles et algorithmes markoviens*
- Marius Iosifescu, *Finite Markov processes and their applications*, John Wiley and sons

Biblio.:

- Olle Häggström, *Finite Markov chains and algorithmic applications*
- Carl Graham, *Chaînes de Markov. Cours et exercices corrigés*, Dunod
- James Norris, *Markov chains*
- Bernard Ycart, *Modèles et algorithmes markoviens*
- Marius Iosifescu, *Finite Markov processes and their applications*, John Wiley and sons

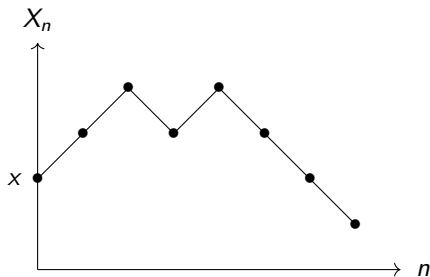
Structure du cours:

- (1) Définition générale, exemples (marches aléatoires)
Dynamique des lois, dynamique des observables, mesures stationnaires
- (2) Réversibilité, spectre de la matrice de transition, convergence vers l'unique mesure invariante
- (3) Problème de Dirichlet, temps d'atteinte
- (4) Algorithme de recuit simulé, algorithmes de Monte Carlo, applications en combinatoire et en physique statistique

Cf. <https://jeremieunterberger.wixsite.com/monsite>, rubrique Enseignement

Exemple 1. Marches aléatoires

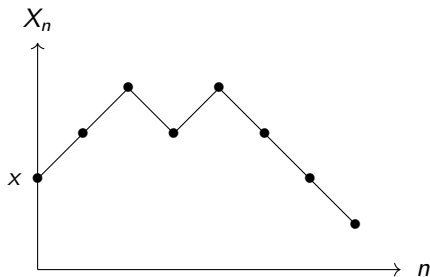
- (dimension 1) $Z_i, i = 1, 2, \dots$ variables de **Rademacher**
 $\mathbb{P}[Z_i = \pm 1] = \frac{1}{2}$
- Position initiale ($t = 0$): $X_0 = x \in \mathbb{Z}$
- Position à l'instant n : $X_n = x + Z_1 + \dots + Z_n$



$X = (X_n)_{n \geq 0}$ v.a. dans $(\Omega, \mathcal{F}, \mathbb{P})$, définissent un **processus stochastique**
 $n \mapsto X_n(\omega)$ **trajectoire aléatoire**

Exemple 1. Marches aléatoires

- (dimension 1) $Z_i, i = 1, 2, \dots$ variables de **Rademacher**
 $\mathbb{P}[Z_i = \pm 1] = \frac{1}{2}$
- Position initiale ($t = 0$): $X_0 = x \in \mathbb{Z}$
- Position à l'instant n : $X_n = x + Z_1 + \dots + Z_n$



$X = (X_n)_{n \geq 0}$ v.a. dans $(\Omega, \mathcal{F}, \mathbb{P})$, définissent un **processus stochastique**
 $n \mapsto X_n(\omega)$ **trajectoire aléatoire**

Loi conditionnelle de la position à $t = n$ en fonction du passé:

$$\mathbb{P}[X_n = y | X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}] = \mathbb{P}[Z_n = y - x_{n-1}]$$

Rappel: $\mathbb{P}[X_n = y | X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}] = \mathbb{P}[Z_n = y - x_{n-1}]$

Propriété de Markov: le processus ne dépend que de son passé immédiat

$\mathbb{P}[X_n = y | X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}]$ ne dépend que de x_{n-1}

La position à l'instant n ne dépend du passé
que via la position à l'instant $n - 1$

Formule des probabilités totales $\Longleftrightarrow$

$$\mathbb{P}[X_n = y | X_0 = x_0, \dots, X_{n-1} = x_{n-1}] = \mathbb{P}[X_n = y | X_{n-1} = x_{n-1}]$$

$\nearrow$
= proba de transition de x_{n-1} vers $y = p_{x_{n-1} \rightarrow y}$

1. Chaîne de Markov: définition générale

$(\Omega, \mathcal{F}, \mathbb{P})$ espace de probabilités

S = espace des états ("space state") supposé fini ou dénombrable

$X = (X_n)_{n \geq 0}$ suite v.a. à valeurs dans S

X est une chaîne de Markov si et seulement si

$$\mathbb{P}[X_n = y | X_0 = x_0, \dots, X_{n-1} = x_{n-1}] = \mathbb{P}[X_n = y | X_{n-1} = x_{n-1}]$$

► Chaîne de Markov homogène:

probas de transition

$$\mathbb{P}[X_n = y | X_{n-1} = x] = p_{x \rightarrow y} \text{ indépendantes de } n$$

loi de transition:

$$\mathbb{P}[X_n = \cdot | X_{n-1} = x] \text{ loi de } X_n \text{ connaissant } X_{n-1} \implies \sum_y p_{x \rightarrow y} = 1$$

Sauf mention contraire, toutes les chaînes considérées dans la suite sont supposées homogènes.

Esquisses d'exemples: modèles physiques, chimiques et biologiques

► Reproduction d'une population homogène

X_n = # nombre d'individus au temps n

loi de transition donnée

Modèle jouet (toy model):

processus de vie et de mort (birth and death process)

$$p_{n \rightarrow n+1} + p_{n \rightarrow n-1} = 1$$

► Génétique et transmission des caractères: reproduction sexuée

X_n = # pois à fleurs blanches (récessif), Y_n = # pois à fleurs rouges (dominant)

aléa: choix de l'allèle lors de la reproduction

Lois de Mendel

► Cinétique chimique



→ cinétique déterministe:

$$\frac{d[C]}{dt} = k[A][B], \quad \frac{d[A]}{dt} = \frac{d[B]}{dt} = -k[A][B]$$

[.] = concentration

► Cinétique chimique

Réaction: $A + B \xrightarrow{k} C$

→ cinétique déterministe:

$$\frac{d[C]}{dt} = k[A][B], \quad \frac{d[A]}{dt} = \frac{d[B]}{dt} = -k[A][B]$$

[.] = concentration

induite par chaîne de Markov $X_n = (a_n, b_n, c_n) \in \mathbb{N}^3$

$$P_{(a,b,c) \rightarrow (a-1,b-1,c+1)} \propto k \times a \times b$$

↖
:= Cste ×

► Cinétique chimique

Réaction: $A + B \xrightarrow{k} C$

→ cinétique déterministe:

$$\frac{d[C]}{dt} = k[A][B], \quad \frac{d[A]}{dt} = \frac{d[B]}{dt} = -k[A][B]$$

$[\cdot]$ =concentration

induite par chaîne de Markov $X_n = (a_n, b_n, c_n) \in \mathbb{N}^3$

$$P_{(a,b,c) \rightarrow (a-1,b-1,c+1)} \propto k \times a \times b$$

↖
:= Cste ×

en ce sens que

$$\mathbb{E}[c_{n+1} | X_n = (a, b, c)] = c + \text{Cste} \times k \times a \times b$$

Cinétique chimique (suite) : trois remarques

- ▶ chaînes de Markov en temps continu: temps de transitions eux-mêmes aléatoires (loi $\text{Exp}(k)$, k =taux de transition)
non étudiées dans ce cours

Cinétique chimique (suite) : trois remarques

► chaînes de Markov en temps continu: temps de transitions eux-mêmes aléatoires
(loi $\text{Exp}(k)$, k =taux de transition)
non étudiées dans ce cours

► version markovienne $\rightarrow$ cinétique déterministe + aléa
théorème central limite: en temps long $N \rightarrow \infty$, $c_N \approx N[C] + O(\sqrt{N})$

↑
fluctuations

temps continu : $t = n/N$, $N \rightarrow \infty \implies \frac{1}{N}c_n \approx [C](t)$

Cinétique chimique (suite) : trois remarques

► chaînes de Markov en temps continu: temps de transitions eux-mêmes aléatoires
(loi $\text{Exp}(k)$, k =taux de transition)
non étudiées dans ce cours

► version markovienne $\rightarrow$ cinétique déterministe + aléa

théorème central limite: en temps long $N \rightarrow \infty$, $c_N \approx N[C] + O(\sqrt{N})$

↑
fluctuations

temps continu : $t = n/N$, $N \rightarrow \infty \implies \frac{1}{N}c_n \approx [C](t)$

► Pourquoi introduire un aléa ?

► fluctuations thermiques (Einstein: cf infra, mouvement brownien)

température=désordre=aléa (Boltzmann) induisant un comportement

irréversible

Cinétique chimique (suite) : trois remarques

► chaînes de Markov en temps continu: temps de transitions eux-mêmes aléatoires (loi $\text{Exp}(k)$, k =taux de transition)
non étudiées dans ce cours

► version markovienne $\rightarrow$ cinétique déterministe + aléa
théorème central limite: en temps long $N \rightarrow \infty$, $c_N \approx N[C] + O(\sqrt{N})$

↑
fluctuations

temps continu : $t = n/N$, $N \rightarrow \infty \implies \frac{1}{N}c_n \approx [C](t)$

► Pourquoi introduire un aléa ?

► fluctuations thermiques (Einstein: cf infra, mouvement brownien)
température=désordre=aléa (Boltzmann) induisant un comportement

irréversible

► fluctuations quantiques (Fermi's golden rule)

‣ Comment passe-t-on de lois fondamentales hamiltoniennes réversibles à un comportement macroscopique irréversible ‣

2. Marches aléatoires

Rappel: $p_{x \rightarrow y} = \mathbb{P}[X_n = y | X_{n-1} = x]$

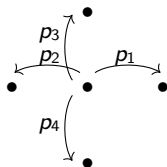
Suite de l'exemple 1, dimension $d = 1, 2, 3, \dots$

$x, y \in \mathbb{Z}^d$ réseau carré

$$p_{x \rightarrow y} = p_{y-x}, \sum_{z \in \mathbb{Z}^d} p_z = 1$$

► Marches aléatoires simples

$$p_z = 0 \text{ sauf si } \|z\|_1 := \sum_{i=1}^d |z_i| = 1$$



$$p_1 = p_{(1,0)}, p_2 = p_{(-1,0)}$$

$$p_3 = p_{(0,1)}, p_4 = p_{(0,-1)}$$

► ► Marches aléatoires simples symétriques

$$p_z = p_{-z} \implies \mathbb{E}[X_{n+1} | X_n = x] = x$$

$$\text{Dém.: } \mathbb{E}[X_{n+1} | X_n = x] = \sum_z p_z(x+z) = x.$$

3. Dynamique de la loi

$\mu_n := \mathbb{P}_{X_n}$ loi du processus au temps n ,

$\mu_0 = \mathbb{P}_{X_0}$ mesure (loi, distribution, condition) initiale

Formule des probabilités totales

$$\begin{aligned}\mu_{n+1}(y) &= \mathbb{P}[X_{n+1} = y] = \sum_x \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] \\ &= \sum_x \mu_n(x) p_{x \rightarrow y} \quad (*)\end{aligned}$$

Matrice de transition

$P := (p_{x \rightarrow y})_{x,y}$ matrice (in)finie

► Cas de la marche aléatoire simple symétrique sur $\mathbb{Z}$:

$$p_{x \rightarrow x \pm 1} = p_{\pm 1} = \frac{1}{2}$$

$$\mu_{n+1}(y) = \frac{1}{2}(\mu_n(y-1) + \mu_n(y+1))$$

$$P = \begin{pmatrix} \ddots & \ddots & & & \\ & \ddots & & & \\ & & 0 & \frac{1}{2} & \\ & & \frac{1}{2} & \ddots & \ddots \\ & & & \ddots & \ddots \end{pmatrix}$$

3. Dynamique de la loi

$\mu_n := \mathbb{P}_{X_n}$ loi du processus au temps n ,

$\mu_0 = \mathbb{P}_{X_0}$ mesure (loi, distribution, condition) initiale

Formule des probabilités totales

$$\begin{aligned}\mu_{n+1}(y) &= \mathbb{P}[X_{n+1} = y] = \sum_x \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] \\ &= \sum_x \mu_n(x) p_{x \rightarrow y} \quad (*)\end{aligned}$$

Matrice de transition

$P := (p_{x \rightarrow y})_{x,y}$ matrice (in)finie

► Cas de la marche aléatoire simple symétrique sur $\mathbb{Z}$:

$$p_{x \rightarrow x \pm 1} = p_{\pm 1} = \frac{1}{2}$$

$$\mu_{n+1}(y) = \frac{1}{2}(\mu_n(y-1) + \mu_n(y+1))$$

$$P = \begin{pmatrix} \ddots & \ddots & & & \\ & \ddots & & & \\ & & 0 & \frac{1}{2} & \\ & & \frac{1}{2} & \ddots & \ddots \\ & & & \ddots & \ddots \end{pmatrix}$$

$$(*) \iff \mu_{n+1} = \mu_n P \quad (\text{Equation de Chapman-Kolmogorov})$$

Mesures μ_n, μ_{n+1} = vecteurs ligne

$\mu_n P$ = vecteur ligne $\times$ matrice carrée = vecteur ligne = mesure

3. Dynamique de la loi (suite)

Si μ =mesure de proba, alors μP aussi:

$$\left(\sum_x \mu_x = 1, \sum_y p_{x \rightarrow y} = 1\right) \implies \left(\sum_y \left(\sum_x \mu_x p_{x \rightarrow y}\right) = 1\right)$$

La suite $n \mapsto \mu_n = \mathbb{P}_{X_n}$ s'appelle dynamique de la loi.

3. Dynamique de la loi (suite)

Si μ =mesure de proba, alors μP aussi:

$$\left(\sum_x \mu_x = 1, \sum_y p_{x \rightarrow y} = 1\right) \implies \left(\sum_y \left(\sum_x \mu_x p_{x \rightarrow y}\right) = 1\right)$$

La suite $n \mapsto \mu_n = \mathbb{P}_{X_n}$ s'appelle **dynamique de la loi**.

► C'est un système dynamique déterministe **sur l'ensemble des mesures de proba**

Chapman-Kolmogorov $\implies \boxed{\mu_n = \mu_0 P^n}$.

3. Dynamique de la loi (suite)

Si μ =mesure de proba, alors μP aussi:

$$\left(\sum_x \mu_x = 1, \sum_y p_{x \rightarrow y} = 1\right) \implies \left(\sum_y \left(\sum_x \mu_x p_{x \rightarrow y}\right) = 1\right)$$

La suite $n \mapsto \mu_n = \mathbb{P}_{X_n}$ s'appelle dynamique de la loi.

- C'est un système dynamique déterministe sur l'ensemble des mesures de proba
Chapman-Kolmogorov $\implies \mu_n = \mu_0 P^n$.
- Mais la suite $n \mapsto X_n$ est un système dynamique aléatoire.

3. Dynamique de la loi (suite)

Si μ = mesure de proba, alors μP aussi:

$$\left(\sum_x \mu_x = 1, \sum_y p_{x \rightarrow y} = 1 \right) \implies \left(\sum_y \left(\sum_x \mu_x p_{x \rightarrow y} \right) = 1 \right)$$

La suite $n \mapsto \mu_n = \mathbb{P}_{X_n}$ s'appelle dynamique de la loi.

- C'est un système dynamique déterministe sur l'ensemble des mesures de proba

Chapman-Kolmogorov $\implies \mu_n = \mu_0 P^n$.

- Mais la suite $n \mapsto X_n$ est un système dynamique aléatoire.

- Cas particulier (chaîne de Markov inhomogène):

$(f_n)_n, f_n : S \rightarrow S$ fonctions

$$\left\{ \begin{array}{l} X_0 = x_0 \text{ déterministe} \\ \mathbb{P}[X_{n+1} = y | X_n = x] = \delta_{y, f_n(x)} \end{array} \right\} \implies$$

système dynamique déterministe $(x_n)_n, x_{n+1} = f_n(x_n)$

3. Dynamique de la loi (suite)

Si μ = mesure de proba, alors μP aussi:

$$\left(\sum_x \mu_x = 1, \sum_y p_{x \rightarrow y} = 1 \right) \implies \left(\sum_y \left(\sum_x \mu_x p_{x \rightarrow y} \right) = 1 \right)$$

La suite $n \mapsto \mu_n = \mathbb{P}_{X_n}$ s'appelle **dynamique de la loi**.

► C'est un système dynamique déterministe **sur l'ensemble des mesures de proba**

Chapman-Kolmogorov $\implies \boxed{\mu_n = \mu_0 P^n}$.

► Mais la suite $n \mapsto X_n$ est un **système dynamique aléatoire**.

► Cas particulier (chaîne de Markov inhomogène):

$(f_n)_n, f_n : S \rightarrow S$ fonctions

$$\left\{ \begin{array}{l} X_0 = x_0 \text{ déterministe} \\ \mathbb{P}[X_{n+1} = y | X_n = x] = \delta_{y, f_n(x)} \end{array} \right\} \implies$$

système dynamique déterministe $(x_n)_n, x_{n+1} = f_n(x_n)$

‡ Un système dynamique aléatoire est-il plus *compliqué* qu'un système dynamique déterministe ‡

3. Dynamique de la loi (suite et fin)

Rappel.

$$\blacktriangleright \mu_{n+1}(y) = \mathbb{P}[X_{n+1} = y] = \sum_x \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] = \sum_x \mu_n(x) p_{x \rightarrow y}$$

3. Dynamique de la loi (suite et fin)

Rappel.

$$\blacktriangleright \mu_{n+1}(y) = \mathbb{P}[X_{n+1} = y] = \sum_x \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] = \sum_x \mu_n(x) p_{x \rightarrow y}$$

Formules duales: dynamique des observables

Observable = fonction sur S

$$\mathbf{f} \in \mathbb{C}^S \longleftrightarrow f : S \rightarrow \mathbb{C}, \quad f(x) = f_x$$

3. Dynamique de la loi (suite et fin)

Rappel.

$$\blacktriangleright \mu_{n+1}(y) = \mathbb{P}[X_{n+1} = y] = \sum_x \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] = \sum_x \mu_n(x) p_{x \rightarrow y}$$

Formules duales: dynamique des observables

Observable=fonction sur S

$$f \in \mathbb{C}^S \longleftrightarrow f : S \rightarrow \mathbb{C}, \quad f(x) = f_x$$

Dynamique des observables: $n \mapsto f_n$, $f_n : x \mapsto \mathbb{E}_x[f(X_n)] = \mathbb{E}[f(X_n^x)]$,

X_n^x =état à l'instant n de la chaîne issue de $x \in S$

$\mathbb{E}_x$ = loi de la chaîne de condition initiale δ_x déterministe

$f_n(x)$ =moyenne des valeurs de f en les états
à l'instant n des trajectoires aléatoires issues de x

3. Dynamique de la loi (suite et fin)

Rappel.

$$\blacktriangleright \mu_{n+1}(y) = \mathbb{P}[X_{n+1} = y] = \sum_x \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] = \sum_x \mu_n(x) p_{x \rightarrow y}$$

Formules duales: dynamique des observables

Observable=fonction sur S

$$f \in \mathbb{C}^S \longleftrightarrow f : S \rightarrow \mathbb{C}, \quad f(x) = f_x$$

Dynamique des observables: $n \mapsto f_n$, $f_n : x \mapsto \mathbb{E}_x[f(X_n)] = \mathbb{E}[f(X_n^x)]$,

$\left| \begin{array}{l} X_n^x = \text{état à l'instant } n \text{ de la chaîne issue de } x \in S \\ \mathbb{E}_x = \text{loi de la chaîne de condition initiale } \delta_x \text{ déterministe} \end{array} \right.$

$f_n(x)$ = moyenne des valeurs de f en les états
à l'instant n des trajectoires aléatoires issues de x

Définition. Pf = matrice carrée $\times$ vecteur colonne = vecteur colonne

$$\longleftrightarrow Pf : x \mapsto (Pf)(x) = \sum_y p_{x \rightarrow y} f(y)$$

Analyse au premier pas: on conditionne suivant les valeurs de X_1 ,

$$\blacktriangleright f_{n+1}(x) = \mathbb{E}[f(X_{n+1}^x)] = \sum_y \mathbb{E}[f(X_{n+1}^x) | X_1^x = y] \mathbb{P}[X_1^x = y] = \sum_y f_n(y) p_{x \rightarrow y}$$

Propriété de Markov

Rappel.

$$f_{n+1}(x) = \mathbb{E}[f(X_{n+1}^x)] = \sum_y \mathbb{E}[f(X_{n+1}^x) | X_1^x = y] \mathbb{P}[X_1^x = y] = \sum_y f_n(y) p_{x \rightarrow y} \quad (*)$$

Propriété de Markov

$$\text{soit } \mathbf{f}_{n+1} = P\mathbf{f}_n \implies \boxed{f_{n+1} = Pf_n, f_n = P^n f}$$

formules duales de $\mu_n = \mu_0 P^n$

Rappel.

$$f_{n+1}(x) = \mathbb{E}[f(X_{n+1}^x)] = \sum_y \mathbb{E}[f(X_{n+1}^x) | X_1^x = y] \mathbb{P}[X_1^x = y] = \sum_y f_n(y) p_{x \rightarrow y} \quad (*)$$

Propriété de Markov

$$\text{soit } f_{n+1} = P f_n \implies \boxed{f_{n+1} = P f_n, f_n = P^n f}$$

formules duales de $\mu_n = \mu_0 P^n$

Propriété de Markov (simple)

$$\begin{aligned} & \mathbb{P}[(X_{m+n}^x)_{n \in \mathbb{N}^*} \in \mathcal{B}; X_m^x = y; (X_p^x)_{0 \leq p \leq m-1} \in \mathcal{A}] \\ &= \mathbb{P}[(X_n^y)_{n \in \mathbb{N}^*} \in \mathcal{B}] \times \mathbb{P}[X_m^x = y; (X_p^x)_{0 \leq p \leq m-1} \in \mathcal{A}] \end{aligned}$$

Appl.: $\mathcal{B} = \{z\}$ ($z \in S$), $m = 1$, $\mathcal{A} = \Omega$, sommer $\sum_z f(z)(\dots) \Rightarrow (*)$

Dém. élémentaire: lois fini-dimensionnelles caractérisent la chaîne $\Rightarrow n \leq N$; $\mathcal{A}, \mathcal{B}$ = singletons

Rappel.

$$f_{n+1}(x) = \mathbb{E}[f(X_{n+1}^x)] = \sum_y \mathbb{E}[f(X_{n+1}^x) | X_1^x = y] \mathbb{P}[X_1^x = y] = \sum_y f_n(y) p_{x \rightarrow y} \quad (*)$$

Propriété de Markov

$$\text{soit } f_{n+1} = P f_n \implies f_{n+1} = P f_n, f_n = P^n f$$

formules duales de $\mu_n = \mu_0 P^n$

Propriété de Markov (simple)

$$\begin{aligned} & \mathbb{P}[(X_{m+n}^x)_{n \in \mathbb{N}^*} \in \mathcal{B}; X_m^x = y; (X_p^x)_{0 \leq p \leq m-1} \in \mathcal{A}] \\ &= \mathbb{P}[(X_n^y)_{n \in \mathbb{N}^*} \in \mathcal{B}] \times \mathbb{P}[X_m^x = y; (X_p^x)_{0 \leq p \leq m-1} \in \mathcal{A}] \end{aligned}$$

Appl.: $\mathcal{B} = \{z\}$ ($z \in S$), $m = 1$, $\mathcal{A} = \Omega$, sommer $\sum_z f(z)(\dots) \Rightarrow (*)$

Dém. élémentaire: lois fini-dimensionnelles caractérisent la chaîne $\Rightarrow n \leq N$; $\mathcal{A}, \mathcal{B}$ = singletons

Généralisation immédiate:

$$\begin{aligned} & \mathbb{P}_{\mu_0}[(X_{m+n})_{n \in \mathbb{N}} \in \mathcal{B}; X_m = y; (X_p)_{0 \leq p \leq m-1} \in \mathcal{A}] \\ &= \mathbb{P}_y[(X_n)_{n \in \mathbb{N}} \in \mathcal{B}] \times \mathbb{P}_{\mu_0}[X_m = y; (X_p)_{0 \leq p \leq m-1} \in \mathcal{A}] \end{aligned}$$

4. Mesures stationnaires

Définition. Une mesure μ sur $(\Omega, \mathcal{F})$ est **stationnaire** pour la (chaîne de Markov de) matrice de transition P si $\mu = \mu P$.

 Lorsque S est infini, bien distinguer: mesures *de proba* stationnaires/mesures stationnaires !

Notation. $\mathcal{M}_+(S) = \{\text{mesures}\}$, $\mathcal{M}_1(S) = \{\text{mesures de proba}\}$, $\mathcal{M}(S) \simeq \mathbb{R}^S$ espace vectoriel des mesures *signées* (=vecteurs lignes)

4. Mesures stationnaires

Définition. Une mesure μ sur $(\Omega, \mathcal{F})$ est **stationnaire** pour la (chaîne de Markov de) matrice de transition P si $\mu = \mu P$.



Lorsque S est infini, bien distinguer: mesures *de proba* stationnaires/mesures stationnaires !

Notation. $\mathcal{M}_+(S) = \{\text{mesures}\}$, $\mathcal{M}_1(S) = \{\text{mesures de proba}\}$, $\mathcal{M}(S) \simeq \mathbb{R}^S$ espace vectoriel des mesures *signées* (=vecteurs lignes)

Ex.: marche aléatoires symétrique simple sur $\mathbb{Z}^d$:

$\mu = C \sum_{x \in \mathbb{Z}^d} \delta_x$ ($C > 0$) mesure stationnaire, mais *pas mesure de proba*

► Si S est fini et $\mu = \mu P$ *mesure*, alors $\tilde{\mu} = \frac{1}{\|\mu\|} \mu$, $\|\mu\| = \mu(S) = \sum_x \mu_x$ est une mesure de proba

4. Mesures stationnaires

Définition. Une mesure μ sur $(\Omega, \mathcal{F})$ est **stationnaire** pour la (chaîne de Markov de) matrice de transition P si $\mu = \mu P$.



Lorsque S est infini, bien distinguer: mesures *de proba* stationnaires/mesures stationnaires !

Notation. $\mathcal{M}_+(S) = \{\text{mesures}\}$, $\mathcal{M}_1(S) = \{\text{mesures de proba}\}$, $\mathcal{M}(S) \simeq \mathbb{R}^S$ espace vectoriel des mesures *signées* (=vecteurs lignes)

Ex.: marche aléatoires symétrique simple sur $\mathbb{Z}^d$:

$\mu = C \sum_{x \in \mathbb{Z}^d} \delta_x$ ($C > 0$) mesure stationnaire, mais *pas mesure de proba*

► Si S est fini et $\mu = \mu P$ *mesure*, alors $\tilde{\mu} = \frac{1}{\|\mu\|} \mu$, $\|\mu\| = \mu(S) = \sum_x \mu_x$ est une mesure de proba

► Calcul pour $d = 1$: $\mu_x = \frac{1}{2}(\mu_{x-1} + \mu_{x+1})$

4. Mesures (de proba) stationnaires (suite)

Rappel: marche aléatoire symétrique simple sur $\mathbb{Z}$: $\mu_x = \frac{1}{2}(\mu_{x-1} + \mu_{x+1})$

4. Mesures (de proba) stationnaires (suite)

Rappel: marche aléatoire symétrique simple sur $\mathbb{Z}$: $\mu_x = \frac{1}{2}(\mu_{x-1} + \mu_{x+1})$

Rappel: suites récurrentes d'ordre 2: $z_{n+1} = az_n + bz_{n-1}$

Equation caractéristique: $r^2 = ar + b$

*Solution générale:
$$\begin{cases} C_1 r_1^n + C_2 r_2^n & (\Delta \neq 0) \\ r^n (C_1 + C_2 n) & (\Delta = 0, \text{ cas résonant}) \end{cases}$$*

► *Exo: $\mu \neq 0$ mesure stationnaire $\implies \mu$ non normalisable, i.e. $\mu(S) = \infty$*

4. Mesures (de proba) stationnaires (suite)

Rappel: marche aléatoire symétrique simple sur $\mathbb{Z}$: $\mu_x = \frac{1}{2}(\mu_{x-1} + \mu_{x+1})$

Rappel: suites récurrentes d'ordre 2: $z_{n+1} = az_n + bz_{n-1}$

Equation caractéristique: $r^2 = ar + b$

*Solution générale:
$$\begin{cases} C_1 r_1^n + C_2 r_2^n & (\Delta \neq 0) \\ r^n (C_1 + C_2 n) & (\Delta = 0, \text{ cas résonant}) \end{cases}$$*

► Exo: $\mu \neq 0$ mesure stationnaire $\implies \mu$ non normalisable, i.e. $\mu(S) = \infty$

► Exo. (1ère variante, cf T.D.) marche aléatoire symétrique simple sur $\mathbb{Z}/N\mathbb{Z}$:

mesures stationnaires $\propto \sum_{x \in \mathbb{Z}/N\mathbb{Z}} \delta_x$ normalisables

4. Mesures (de proba) stationnaires (suite)

Rappel: marche aléatoire symétrique simple sur $\mathbb{Z}$: $\mu_x = \frac{1}{2}(\mu_{x-1} + \mu_{x+1})$

Rappel: suites récurrentes d'ordre 2: $z_{n+1} = az_n + bz_{n-1}$

Equation caractéristique: $r^2 = ar + b$

*Solution générale:
$$\begin{cases} C_1 r_1^n + C_2 r_2^n & (\Delta \neq 0) \\ r^n (C_1 + C_2 n) & (\Delta = 0, \text{ cas résonant}) \end{cases}$$*

► Exo: $\mu \neq 0$ mesure stationnaire $\implies \mu$ non normalisable, i.e. $\mu(S) = \infty$

► Exo. (1ère variante, cf T.D.) marche aléatoire symétrique simple sur $\mathbb{Z}/N\mathbb{Z}$:

mesures stationnaires $\propto \sum_{x \in \mathbb{Z}/N\mathbb{Z}} \delta_x$ normalisables

► Exo. (autres variantes, cf T.D.) conditions d'absorption/de réflexion au bord $\longrightarrow$ mesures inhomogènes

4. Mesures (de proba) stationnaires (suite)

Rappel: marche aléatoire symétrique simple sur $\mathbb{Z}$: $\mu_x = \frac{1}{2}(\mu_{x-1} + \mu_{x+1})$

Rappel: suites récurrentes d'ordre 2: $z_{n+1} = az_n + bz_{n-1}$

Equation caractéristique: $r^2 = ar + b$

*Solution générale:
$$\begin{cases} C_1 r_1^n + C_2 r_2^n & (\Delta \neq 0) \\ r^n (C_1 + C_2 n) & (\Delta = 0, \text{ cas résonant}) \end{cases}$$*

► Exo: $\mu \neq 0$ mesure stationnaire $\implies \mu$ non normalisable, i.e. $\mu(S) = \infty$

► Exo. (1ère variante, cf T.D.) marche aléatoire symétrique simple sur $\mathbb{Z}/N\mathbb{Z}$:

mesures stationnaires $\propto \sum_{x \in \mathbb{Z}/N\mathbb{Z}} \delta_x$ normalisables

► Exo. (autres variantes, cf T.D.) conditions d'absorption/de réflexion au bord $\longrightarrow$ mesures inhomogènes

► Exo. (autres variantes, cf T.D.) cas des marches aléatoires simples asymétriques



Forte sensibilité aux conditions au bord !

$\longrightarrow$? Limite en volume infini ?

4. Mesures (de proba) stationnaires (suite)

Complément. Rappel: $S = \{x_n\}$ fini ou dénombrable

Notation (rappel): $\mathcal{M}_1(S) := \{\mu : S \rightarrow \mathbb{R}_+ \mid \mu(S) = 1\}$, $\mathcal{M}(S) = \{\text{mesures signées}\}$

Distance en variation totale:

$\mu, \nu \in \mathcal{M}_1(S) \longrightarrow \mu - \nu : S \rightarrow \mathbb{R}$ mesure signée

$$d(\mu, \nu) := |\mu - \nu| \equiv \frac{1}{2} \sum_x |\mu_x - \nu_x| \in [0, 1]$$

$\implies \mathcal{M}_1(S)$ espace métrique borné

|| Si S fini alors $\mathcal{M}_1(S) \subset \mathbb{R}^S$ compact

|| $\rightarrow$ * THEOREME EXISTENCE MESURES STATIONNAIRES *

4. Mesures (de proba) stationnaires (suite)

Complément. Rappel: $S = \{x_n\}$ fini ou dénombrable

Notation (rappel): $\mathcal{M}_1(S) := \{\mu : S \rightarrow \mathbb{R}_+ \mid \mu(S) = 1\}$, $\mathcal{M}(S) = \{\text{mesures signées}\}$

Distance en variation totale:

$\mu, \nu \in \mathcal{M}_1(S) \rightarrow \mu - \nu : S \rightarrow \mathbb{R}$ mesure signée

$$d(\mu, \nu) := |\mu - \nu| \equiv \frac{1}{2} \sum_x |\mu_x - \nu_x| \in [0, 1]$$

$\Rightarrow \mathcal{M}_1(S)$ espace métrique borné

|| Si S fini alors $\mathcal{M}_1(S) \subset \mathbb{R}^S$ compact

|| $\rightarrow$ * THEOREME EXISTENCE MESURES STATIONNAIRES *

Lemme. Les matrices de transition sont 1-Lipschitz, i.e. $d(\mu P, \nu P) \leq d(\mu, \nu)$.

Preuve.

$$2d(\mu P, \nu P) = \sum_y |(\mu P)_y - (\nu P)_y| = \sum_y \left| \sum_x (\mu_x - \nu_x) p_{x \rightarrow y} \right| \leq \sum_x |\mu_x - \nu_x| \left(\sum_y p_{x \rightarrow y} \right) \overset{1}{=} \sum_x |\mu_x - \nu_x|$$

□

4. Mesures (de proba) stationnaires (suite)

Complément. Rappel: $S = \{x_n\}$ fini ou dénombrable

Notation (rappel): $\mathcal{M}_1(S) := \{\mu : S \rightarrow \mathbb{R}_+ \mid \mu(S) = 1\}$, $\mathcal{M}(S) = \{\text{mesures signées}\}$

Distance en variation totale:

$\mu, \nu \in \mathcal{M}_1(S) \rightarrow \mu - \nu : S \rightarrow \mathbb{R}$ mesure signée

$$d(\mu, \nu) := |\mu - \nu| \equiv \frac{1}{2} \sum_x |\mu_x - \nu_x| \in [0, 1]$$

$\Rightarrow \mathcal{M}_1(S)$ espace métrique borné

|| Si S fini alors $\mathcal{M}_1(S) \subset \mathbb{R}^S$ compact

|| $\rightarrow$ * THEOREME EXISTENCE MESURES STATIONNAIRES *

Lemme. Les matrices de transition sont 1-Lipschitz, i.e. $d(\mu P, \nu P) \leq d(\mu, \nu)$.

Preuve.

$$2d(\mu P, \nu P) = \sum_y |(\mu P)_y - (\nu P)_y| = \sum_y \left| \sum_x (\mu_x - \nu_x) p_{x \rightarrow y} \right| \leq \sum_x |\mu_x - \nu_x| \left(\sum_y p_{x \rightarrow y} \right) = \sum_x |\mu_x - \nu_x| = 2d(\mu, \nu).$$

□

Corollaire. Si $\mu P^n \xrightarrow{n \rightarrow \infty} \nu$, alors ν est une mesure stationnaire.

Preuve. $\nu_n := \mu P^n \rightarrow \nu$ donc d'après le lemme, $\nu_{n+1} = \nu_n P \rightarrow \nu P$.

□

4. Mesures (de proba) stationnaires (suite)

Complément. Rappel: $S = \{x_n\}$ fini ou dénombrable

Notation (rappel): $\mathcal{M}_1(S) := \{\mu : S \rightarrow \mathbb{R}_+ \mid \mu(S) = 1\}$, $\mathcal{M}(S) = \{\text{mesures signées}\}$

Distance en variation totale:

$\mu, \nu \in \mathcal{M}_1(S) \rightarrow \mu - \nu : S \rightarrow \mathbb{R}$ mesure signée

$$d(\mu, \nu) := |\mu - \nu| \equiv \frac{1}{2} \sum_x |\mu_x - \nu_x| \in [0, 1]$$

$\Rightarrow \mathcal{M}_1(S)$ espace métrique borné

|| Si S fini alors $\mathcal{M}_1(S) \subset \mathbb{R}^S$ compact

|| $\rightarrow$ * THEOREME EXISTENCE MESURES STATIONNAIRES *

Lemme. Les matrices de transition sont 1-Lipschitz, i.e. $d(\mu P, \nu P) \leq d(\mu, \nu)$.

Preuve.

$$2d(\mu P, \nu P) = \sum_y |(\mu P)_y - (\nu P)_y| = \sum_y \left| \sum_x (\mu_x - \nu_x) p_{x \rightarrow y} \right| \leq \sum_x |\mu_x - \nu_x| \left(\sum_y p_{x \rightarrow y} \right) = \sum_x |\mu_x - \nu_x| \overset{1}{=} 1$$

□

Corollaire. Si $\mu P^n \xrightarrow{n \rightarrow \infty} \nu$, alors ν est une mesure stationnaire.

Preuve. $\nu_n := \mu P^n \rightarrow \nu$ donc d'après le lemme, $\nu_{n+1} = \nu_n P \rightarrow \nu P$.

□

|| Les mesures stationnaires s'obtiennent comme limites $\lim_{n \rightarrow \infty} \mu P^n$, μ variant parmi les mesures initiales.

! Comment déterminer les mesures stationnaires !

? Comment déterminer les mesures stationnaires ?

Une classe particulière fréquente et signifiante. Cas des chaînes de Markov réversibles.

Définition. On dit que P est **réversible** s'il existe $(a_x)_{x \in S}$ coeffs ≥ 0 non tous nuls t.q.

$$\forall x, y \in S, \quad a_y p_{y \rightarrow x} = a_x p_{x \rightarrow y}. \quad (\text{Rev})$$

Exo. $\nu_x = \frac{a_x}{\sum_y a_y}$ mesure stationnaire.

► C'est plus simple de chercher des coeffs $(a_x)_x$ vérifiant (Rev) que de résoudre une équation aux valeurs propres.

‡ Comment déterminer les mesures stationnaires ‡

Une classe particulière fréquente et significative. Cas des chaînes de Markov réversibles.

Définition. On dit que P est **réversible** s'il existe $(a_x)_{x \in S}$ coeffs ≥ 0 non tous nuls t.q.

$$\forall x, y \in S, \quad a_y p_{y \rightarrow x} = a_x p_{x \rightarrow y}. \quad (\text{Rev})$$

Exo. $\nu_x = \frac{a_x}{\sum_y a_y}$ mesure stationnaire.

► C'est plus simple de chercher des coeffs $(a_x)_x$ vérifiant (Rev) que de résoudre une équation aux valeurs propres.

Exo. (cf. TD) Si P est réversible, on a une *formule explicite simple* pour les mesures stationnaires (notamment lorsque celle-ci est unique).

S espace d'état fini ou dénombrable

► *Rappel.* $\mathcal{M}_+(S) = \{\text{mesures}\}$, $\mathcal{M}_1(S) = \{\text{mesures de proba}\}$; $\mathcal{M}(S) = \{\text{mesures signées}\}$

► *Rappel.* Si S fini il existe *au moins une* mesure stationnaire (normalisable).

► *Remarque.* Si $\nu_1, \nu_2, \dots$ mesures stationnaires, alors *toute combinaison linéaire à coeffs ≥ 0 aussi.* En particulier, $\forall t \in [0, 1], t\nu_1 + (1 - t)\nu_2 \in \mathcal{M}_1(S)$ si $\nu_1, \nu_2 \in \mathcal{M}_1(S)$.

↙
comb. lin. convexe

Critère simple d'unicité des mesures stationnaires (Théorème de Perron-Frobenius)

Définition. P est **irréductible** si $\forall x, y \in S$, il existe un *chemin* de x vers y ,

$$x \rightarrow x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_n \rightarrow y, \quad p_{x \rightarrow x_1}, p_{x_1 \rightarrow x_2}, \dots, p_{x_n \rightarrow y} > 0$$

S espace d'état fini ou dénombrable

► *Rappel.* $\mathcal{M}_+(S) = \{\text{mesures}\}$, $\mathcal{M}_1(S) = \{\text{mesures de proba}\}$; $\mathcal{M}(S) = \{\text{mesures signées}\}$

► *Rappel.* Si S fini il existe *au moins une* mesure stationnaire (normalisable).

► *Remarque.* Si $\nu_1, \nu_2, \dots$ mesures stationnaires, alors *toute combinaison linéaire à coeffs ≥ 0 aussi*. En particulier, $\forall t \in [0, 1], t\nu_1 + (1 - t)\nu_2 \in \mathcal{M}_1(S)$ si $\nu_1, \nu_2 \in \mathcal{M}_1(S)$.

comb. lin. convexe

Critère simple d'unicité des mesures stationnaires (Théorème de Perron-Frobenius)

Définition. P est **irréductible** si $\forall x, y \in S$, il existe un *chemin* de x vers y ,

$$x \rightarrow x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_n \rightarrow y, \quad p_{x \rightarrow x_1}, p_{x_1 \rightarrow x_2}, \dots, p_{x_n \rightarrow y} > 0$$

Exo. Soit ν mesure t.q. $\nu P = \nu$. Si P est irréductible, alors $\nu > 0$ ($\nu_x > 0$ pour tout x).

Exo. Soit P irréductible. Alors l'espace propre $\{\nu \in \mathcal{M}(S) \mid \nu P = \nu\}$ est de dim. 1, et est engendré par $\nu \in \mathcal{M}_1(S)$ mesure de proba.