Chapitre 6

Chaînes de Markov

6.1 Définition et caractérisations

6.1.1 Définition

Soit S un ensemble fini ou dénombrable, ν une mesure de probabilité sur S et $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice à coefficients positifs. Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires définies sur un espace $(\Omega, \mathcal{F}, \mathbb{P})$. On dit que la suite $(X_n)_{n\geq 0}$ est une *chaîne de Markov* de loi initiale ν et de matrice de passage P si l'on a, pour tout entier $n \geq 1$ et toute suite $x_0, \dots x_n$ d'éléments de S :

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots X_n = x_n) = \nu(x_0) \prod_{i=0}^{n-1} p_{x_i, x_{i+1}}.$$

Exemple: une suite $(X_n)_{n\geq 0}$ de variables aléatoires indépendantes de même loi ν à valeurs dans S dénombrable est une chaîne de Markov. En effet, il suffit de poser pour $(i,j) \in S \times S$ $p_{i,j} = \nu(j)$.

6.1.2 Caractérisation par l'espérance conditionnelle

Théorème 58. *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires à valeurs dans S. Les trois propriétés suivantes sont équivalentes :*

1. *$(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de passage P à valeurs dans S.*
2. *Quels que soient $x_0, \dots, x_{n-1}$ dans S tels que*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots X_{n-1} = x_{n-1}) > 0,$$

alors

$$\mathbb{P}(X_n = x_n | X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = p_{x_{n-1},x_n}.$$

3. $$\mathbb{P}(X_n = x_n | X_0, \dots, X_{n-1}) = p_{X_{n-1},x_n}. \tag{6.1}$$

(Rappelons que $\mathbb{P}(X_n = x_n | X_0, \dots, X_{n-1})$ est une abréviation pour $\mathbb{E}[\mathbb{1}_{\{X_n=x_n\}} | X_0, \dots, X_{n-1}) = \mathbb{E}[\mathbb{1}_{\{X_n=x_n\}} | \sigma(X_0, \dots, X_{n-1})]$.)

Cela signifie que toute l'information que $X_0, \dots, X_{n-1}$ peuvent nous apporter sur X_n est comprise dans X_n.

Démonstration. L'équivalence entre (1) et (2) est immédiate. Pour le reste, on pourra se reporter aux techniques développées dans le corollaire 2. ✓

Remarque: L'équation (6.1) implique que $\mathbb{P}(X_n = x_n | X_{n-1}) = p_{X_{n-1},x_n}$.

6.1.3 Dynamique markovienne

Qu'est ce, concrètement, qu'une chaîne de Markov ? On va voir que c'est une suite de réalisations, au cours du temps, des états d'un système soumis à des transformations aléatoires, la suite des transformations est une suite de transformations indépendantes, de même loi. Évidemment, le résultat de la transformation dépend de la transformation choisie et de l'état du système avant la transformation.

Si $(\Omega, \mathcal{F})$ est un espace mesuré, on appelle "fonction aléatoire" toute application mesurable de $(\Omega, \mathcal{F})$ dans $(S^S, \mathcal{B}(S^S))$.

Comme $\mathcal{B}(S^S)$ est engendrée par les projections sur les coordonnées, $f : \Omega \to S^S = \mathcal{F}(S, S)$ est une fonction aléatoire si et seulement si pour tout $i \in S$, l'application $\omega \mapsto f(\omega)(i)$ est une variable aléatoire. La tribu engendrée par une fonction aléatoire f est la tribu engendrée par les variables $f(\cdot)(i)$, où i décrit S.

Si f est une fonction aléatoire et X une variable aléatoire, $f(X)$ est une variable aléatoire car

$$\{f(X) \in B\} = \bigcup_{i \in S} \{X = i\} \cap \{f(i) \in B\}.$$

Lemme 10. *Soit S un ensemble fini ou dénombrable, ν une loi sur S et χ une mesure sur $(S^S, \mathcal{B}(S^S))$.*

Soit $(f_n)_{n\geq 1}$ une suite de fonctions aléatoires indépendantes de loi χ et X_0 une variable aléatoire de loi μ indépendante de $(f_n)_{n\geq 1}$. On définit $(X_n)_{n\geq 1}$ par

$$\forall n \geq 0 \quad X_{n+1} = f_{n+1}(X_n).$$

Alors $(X_n)_{n\geq 0}$ est une chaîne de Markov de loi initiale ν et de matrice de transition M, où M est définie par

$$\forall (i,j) \in S \times S \quad m_{i,j} = \chi(\{f \in S^S; f(i) = j\}).$$

Démonstration. Soit $A \subset S^{\{0,\dots,n\}}$.

$$\begin{aligned}
&\mathbb{P}(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\} \cap \{X_{n+1} = j\}) \\
&= \mathbb{P}(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\} \cap \{f_{n+1}(i) = j\}) \\
&= \mathbb{P}(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\})\mathbb{P}(f_{n+1}(i) = j) \\
&= \mathbb{P}(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\})\mathbb{P}(f_{n+1} \in S \times \dots \{j\} \times \dots S) \\
&= \mathbb{P}(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\})\chi(S \times \dots \{j\} \times \dots S) \\
&= \mathbb{P}(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\})m_{i,j}.
\end{aligned}$$

✓

Exemple: la marche de l'ivrogne (ou marche aléatoire sur $\mathbb{Z}$)
Un ivrogne sort du café, passablement éméché. À chaque pas, il prend une décision (enfin, si tant est que cela lui soit possible...) : aller à gauche, ou aller à droite. Si on repère par X_n sa position dans la rue

au temps n, on a $S = \mathbb{Z}$, $X_{n+1} = f_{n+1}(X_n)$, où f_n est une suite de translations indépendantes :

$$\mathbb{P}(f_n = (x \mapsto x+1)) = \mathbb{P}(f_n = (x \mapsto x-1)) = 1/2.$$

Comme on le verra un peu plus loin, ce procédé permet de fabriquer toutes les chaînes de Markov.

Pour l'heure, cette notion de fonction aléatoire peut sembler un peu abstraite. En fait, cette fonction aléatoire qui change l'état de la chaîne de Markov est souvent obtenue à partir d'un aléa extérieur, modélisant par exemple l'évolution de l'environnement, et d'un mécanisme déterministe décrivant comment l'état de la chaîne est modifié par l'aléa extérieur.

Corollaire 16. *Soit S un ensemble dénombrable,T un ensemble non-vide. On suppose qu'on dispose d'une fonction $F : S \times T \to T$. Alors, si X_0 est une variable aléatoire à valeurs dans S et $(Y_n)_{n\ge 1}$ une suite de variables aléatoires i.i.d. à valeurs dans T indépendante de X_0, la suite de variables aléatoires $(X_n)_{n\ge 0}$ définie par la récurrence*

$$\forall n \ge 0 \quad X_{n+1} = F(X_n, Y_{n+1})$$

est une chaîne de Markov dont la dynamique est donnée par la matrice $(p_{i,j})_{i,j\in S}$ avec

$$p_{i,j} = \mathbb{P}(F(i, Y_1) = j).$$

Démonstration. Il suffit d'appliquer le corollaire précédent avec la suite de fonctions aléatoires $f_n = F(\cdot, Y_n)$. ✓

Ce point de vue est particulièrement adapté à la simulation sur ordinateur. On notera qu'on n'a pas besoin d'écrire la matrice de la chaîne de Markov; plus on colle de près à la dynamique, plus l'écriture du programme est facile

Exemple: On considère la marche aléatoires biaisée sur le tore $\mathbb{Z}/N\mathbb{Z}$: à chaque pas, on ajoute 1 avec probabilité p, et on enlève 1 avec probabilité $1-p$.

On propose le code Julia suivant, qui retourne un vecteur représentant une trajectoire entre les temps 0 et $t_{\max}$, pour une chaîne de Markov partant de x_0, et suivant la dynamique déterminée par les paramètres N et p.

```
function marche_tore(N,p,x0,tmax)
    x=x0 ;  t=[x]
    for i=1:tmax
        if (rand()<p)
            x=mod(x+1,N)
        else
            x=mod(x-1,N)
        end
        t=[t x]
    end
    return(t)
end
```

On notera que le code proposé n'accède pas au tableau via les indices, qui sont une source fréquente d'erreurs de programmation. On peut éviter l'emploi du `if` en notant que si U suit la loi uniforme sur $[0,1]$, $a+(b-a)\mathbb{1}_{\{U\leq p\}} \sim (1-p)\delta_a + p\delta_b$. Cela donne l'écriture alternative :

```
function marche_tore_deux(N,p,x0,tmax)
    x=x0 ;  t=[x]
    for i=1:tmax
        x=mod(x+2*(rand()<p)-1,N)
        t=[t x]
    end
    return(t)
end
```

ou encore

```
function marche_tore_trois(N,p,x0,tmax)
  a=cumsum([x0 2*(rand(1,tmax).<p)-1],2)
  return(mod(a,N))
end
```

Le tout est de trouver un compromis acceptable entre concision et lisibilité !

On peut alors afficher une simulation. [1]

```
  using PyPlot
tmax=50 ; N=11 ; p=3/4
x=collect(0:tmax)
t=marche_tore(N,p,0,tmax)
plot(x,t')
```

On choisit d'afficher les 50 premiers pas d'une marche sur $\mathbb{Z}/11\mathbb{Z}$ partant de 0 et biaisée dans le sens positif ($p = 3/4$).

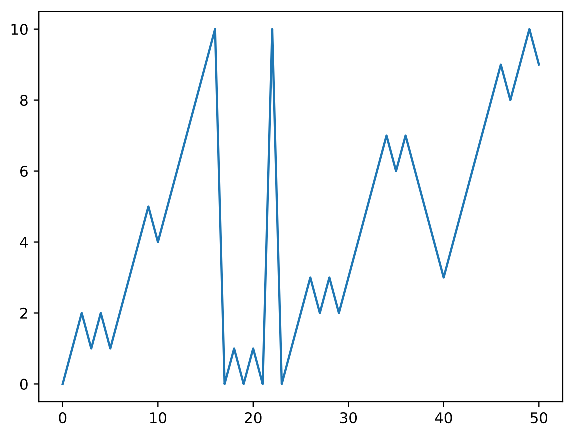


FIGURE 6.1 – Une trajectoire partant de 0 pour $N = 11$, $p = 3/4$. On voit que la courbe a tendance à monter.

1. Pour afficher des graphiques en Julia, il faut installer une bibliothèque appropriée. J'ai choisi ici d'utiliser la librairie PyPlot, qui vient de Python. Pour l'installer, l'ordinateur étant connecté à Internet, il suffit de lancer Julia puis de saisir la commande `Pkg.add("PyPlot")`.

6.2 La modélisation ou les mathématiques ?

Le succès de la théorie des chaînes de Markov a évidemment été déterminé par son efficacité dans la modélisation de phénomènes issus de la vie réelle.

Il faut noter que le choix du modèle n'est pas toujours évident. En particulier, deux modélisations raisonnables d'un même phénomène ne mènent pas nécessairement aux mêmes problèmes mathématiques.

Considérons par exemple le problème suivant : on effectue une suite de lancers d'un dé à 6 faces, et on s'intéresse au nombre de faces différentes qui sont apparues parmi les n premiers lancers.

Une modélisation possible est la suivante : supposons qu'au temps n, k faces différentes sont apparues ; alors parmi les 6 lancers possibles au $n+1$-ème coup, il y en a k qui ne modifient pas le nombre de faces apparues, et $6-k$ qui le font baisser de 1 ; ainsi on peut modéliser le nombre F_n de faces apparues au temps n par une chaîne de Markov déterminée par les transitions

$$p_{i,j} = \begin{cases} \frac{i}{6} & \text{si } i = j \\ 1 - \frac{i}{6} & \text{si } j = i+1 \\ 0 & \text{sinon} \end{cases} .$$

Ici, la modélisation est raisonnable, mais pas non plus complètement évidente. Le choix des coefficients n'est pas en cause, puisqu'il résulte d'un dénombrement facile ; ce qui pose question, c'est le caractère markovien de la suite (F_n), qui n'est pas évident. Ici, la markovianité est posée en hypothèse ; elle n'est pas démontrée par un raisonnement mathématique.

Une autre possibilité est de « remonter » la modélisation. On modélise la suite des lancers par une suite (X_n) de variables aléatoires indépendantes suivant la loi uniforme sur $\{1,2,3,4,5,6\}$, puis, on pose alors $F_n = |\{X_1,\dots,X_n\}|$. La modélisation est ainsi plus proche de la réalité physique, et semble moins arbitraire. Ici, le caractère markovien de (F_n) n'est pas une hypothèse : c'est une conjecture que l'on peut démontrer – ce que nous ferons en exercice. On peut toutefois déjà remarquer que l'ensemble $E_n = \{X_1,\dots,X_n\}$ des faces qui sont été observées au temps n est une chaîne de Markov. En effet, on a la récurrence $E_{n+1} = F(E_n, X_{n+1})$ avec $E_0 = \varnothing$ et

$$F(E, x) = E \cup \{x\},$$

ce qui permet d'appliquer le corollaire 16.

Pour montrer que $Y_n = |E_n|$, il faudra encore un peu travailler, car l'image d'une chaîne de Markov n'est pas toujours une chaîne de Markov (voir les exercices). En particulier, ici, la markovianité de la chaîne est dûe au fait que le dé est supposé équilibré. On verra que si le dé est pipé, $(E_n)_{n\geq 0}$ est toujours une chaîne de Markov, mais pas $(Y_n)_{n\geq 0}$, comme quoi l'hypothèse de markovianité faite dans la première modélisation n'était pas complètement évidente.

6.3 Matrice stochastique

Définition: Soit S un ensemble dénombrable et $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice à coefficients positifs. On dit que P est une matrice stochastique si on a

$$\forall i \in S \quad \sum_{j\in S} p_{i,j} = 1.$$

6.3.1 Existence des chaînes de Markov

Théorème 59. *Soit S un ensemble dénombrable, $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice stochastique et ν une mesure de probabilité sur S. Alors, on peut construire une chaîne de Markov de loi initiale ν et de matrice de passage P.*

Démonstration. Définissons une mesure χ_P sur S^S par

$$\chi_P = \bigotimes_{i\in S} \mu_i,$$

où μ_i est la mesure sur S définie par $\mu_i(j) = p_{i,j}$. Alors χ_P vérifie

$$\chi_P(S \times \dots \{j\} \times \dots S) = p_{i,j}$$

et il suffit d'appliquer le lemme 10. ✓

Lorsque la matrice P est fixée, on note souvent $\mathbb{P}^\nu$ une probabilité sous laquelle $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition P telle que la loi de X_0 sous $\mathbb{P}^\nu$ est ν.[2] De même, on note $\mathbb{E}^\nu$ l'espérance correspondante. Dans le cas où la loi initiale est une masse de Dirac, on écrit simplement $\mathbb{P}^i$ (resp. $\mathbb{E}^i$) au lieu de $\mathbb{P}^{\delta_i}$ (resp. $\mathbb{E}^{\delta_i}$).
Remarque: on est souvent amené à réaliser une telle chaîne sur l'espace canonique $\Omega = S^{\mathbb{N}}$. Dans ce cas, les $(X_k)_{k\geq 0}$ sont les opérateurs de projection canonique : $X_k(\omega) = \omega_k$ et $\mathbb{P}^\nu$ est l'unique mesure sur Ω telle que pour tout entier $n \geq 1$ et toute suite $x_0, \dots x_n$ d'éléments de S :

$$\mathbb{P}^\nu(X_0 = x_0, X_1 = x_1, \dots X_n = x_n) = \nu(x_0) \prod_{i=0}^{n-1} p_{x_i,x_{i+1}}.$$

Corollaire 17. *Soit P une matrice markovienne sur S. Pour tout ν, on note P^ν la mesure markovienne sur $S^{\mathbb{N}}$ de loi initiale ν et de matrice de passage P, ainsi que $\mathbb{P}^i = \mathbb{P}^{\delta_i}$. Pour toute loi ν sur S, $\mathbb{P}^\nu$ admet la désintégration*

$$\mathbb{P}^\nu = \int_S \mathbb{P}^i \, d\nu(i) \tag{6.2}$$

c'est à dire que pour tout borélien A de $S^{\mathbb{N}}$, on a

$$\mathbb{P}^\nu(A) = \int_S \mathbb{P}^i(A) \, d\nu(i) \tag{6.3}$$

Démonstration. Les détails de la preuve ont été donnés dans l'exercice 50 du chapitre 5. On en rappelle les grandes lignes pour soutenir l'intuition : il suffit de définir une mesure μ par

$$\mu(A) = \int_S \mathbb{P}^i(A) \, d\nu(i)$$

et de vérifier que l'on a pour tout entier $n \geq 1$ et toute suite $x_0, \dots x_n$ d'éléments de S :

$$\mu(X_0 = x_0, X_1 = x_1, \dots X_n = x_n) = \nu(x_0) \prod_{i=0}^{n-1} p_{x_i,x_{i+1}}.$$

✓

2. L'existence de la mesure $\mathbb{P}^\nu$ ou de la mesure χ_P relève de la théorie de la mesure. Elle a été établie au chapitre 5.

Remarque: On trouve parfois la notation $\mathbb{P}_i$ à la place de $\mathbb{P}^i$. Dans ce cas, il faut faire attention qu'il peut y avoir ambiguité sur le sens de la notation $\mathbb{P}_X$.

6.3.2 Puissances des matrices stochastiques

Théorème 60. *Soit (X_n) une chaîne de Markov de matrice de transition P et de loi initiale $\mathbb{P}_{X_0} = \nu$. Alors, la loi μ_n de la chaîne au temps n s'écrit $\mu_n = \nu P^n$, où on a écrit ν et μ_n comme des vecteurs lignes.*

Démonstration. Il suffit de montrer que $\mu_{n+1} = \mu_n P$, puis procéder par récurrence sur n. D'après le principe de partition, on a

$$\begin{aligned}
\mu_{n+1}(j) &= \mathbb{P}^\nu(X_{n+1} = j) \\
&= \sum_{i\in S} \mathbb{P}^\nu(X_n = i, X_{n+1} = j) \\
&= \sum_{i\in S} \mathbb{P}^\nu(X_n = i) p_{i,j} \\
&= \sum_{i\in S} \mu_n(i) p_{i,j} \\
&= (\mu_n M)(j).
\end{aligned}$$

✓

En particulier, en prenant $\nu = \delta_i$, on a le corollaire important :

Corollaire 18. *Soit (X_n) une chaîne de Markov à valeur dans S, de matrice de transition P et de loi initiale δ_i, avec $i \in S$. Alors, pour tout $j \in S$, on a*

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

6.3.3 Chaînes de Markov indépendantes

Le résultat qui suit montre comment fabriquer une chaîne de Markov avec deux chaînes de Markov indépendantes. Il sera souvent utilisé comme brique de base pour montrer des résultats plus poussés.

Lemme 11. *Soit $(X_n)_{n\geq 0}$ une chaîne de Markov de matrice de transition M et de loi initiale μ , $(Y_n)_{n\geq 0}$ une chaîne de Markov de matrice de transition N et de loi initiale ν. On suppose en outre que les suites $(X_n)_{n\geq 0}$ et $(Y_n)_{n\geq 0}$ sont indépendantes sous $\mathbb{P}$. Alors la suite $(Z_n)_{n\geq 0}$ définie par $Z_n = (X_n, Y_n)$ est une chaîne de Markov de matrice de transition $M \otimes N$, où $M \otimes N$ est définie par*

$$\forall((i,j),(k,l)) \in S^2 \times S^2 \quad (M \otimes N)(i,j),(k,l) = M(i,k)N(j,l).$$

Démonstration. Soient $(x_0, \dots, x_n) \in S^{n+1}$ et $(y_0, \dots, y_n) \in S^{n+1}$.

$$\begin{aligned}
&\mathbb{P}(\forall i \in \{0,n\}(X_i, Y_i) = (x_i, y_i)) \\
&= \mathbb{P}(\{\forall i \in \{0,n\}X_i = x_i\} \cap \{\forall i \in \{0,n\}Y_i = y_i\}) \\
&= \mathbb{P}(\forall i \in \{0,n\}X_i = x_i)\mathbb{P}(\forall i \in \{0,n\}Y_i = y_i) \\
&= \mu(\{x_0\}) \prod_{i=0}^{n-1} m_{x_i,x_{i+1}} \times \nu(\{y_0\}) \prod_{i=0}^{n-1} n_{y_i,y_{i+1}} \\
&= \mu(\{x_0\})\nu(\{y_0\}) \prod_{i=0}^{n-1} m_{x_i,x_{i+1}} n_{y_i,y_{i+1}} \\
&= (\mu \otimes \nu)(\{x_0, y_0\}) \prod_{i=0}^{n-1} (M \otimes N)((x_i, y_i), (x_{i+1}, y_{i+1})).
\end{aligned}$$

✓

6.3.4 Graphe associé à une matrice stochastique

Soit $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice stochastique. On peut associer à la matrice P (où aux chaînes de Markov correspondantes) un graphe orienté $G = (S, A)$ avec

$$A = \{(x,y) \in S \times S; p_{i,j} > 0\}.$$

Considérons une chaîne de Markov associée à la matrice stochastique P avec la condition initiale déterministe x_0, autrement dit $\nu = \delta_{x_0}$ et notons $\mathbb{P}^{x_0}$ la mesure de probabilité correspondante Alors, comme

$$\mathbb{P}^{x_0}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \prod_{i=0}^{n-1} p_{x_i,x_{i+1}},$$

il est clair que $\mathbb{P}^{x_0}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n)$ est non nul si et seulement si $(x_0, x_1, \dots, x_n)$ constitue un chemin dans le graphe G.

D'après le principe de partition, on a pour une chaîne de Markov avec une loi initiale δ_i :

$$\begin{aligned}&\mathbb{P}^i(X_n = x_n)\\ &= \sum_{(x_0,\dots x_{n-1})\in S^n} \mathbb{P}^i(X_0 = x_0, X_1 = x_1, \dots X_{n-1} = x_{n-1}, X_n = x_n).\end{aligned}$$

En particulier, si l'on pose $p_{i,j}^{(n)} = \mathbb{P}^i(X_n = j)$, on a

$$p_{i,j}^{(n)} = \sum_{x\in S^{n-1}} \mathbb{P}^i(X_1 = x_1, X_2 = X_2, \dots X_{n-1} = x_{n-1}, X_n = j).$$

Donc $p_{i,j}^{(n)} > 0$, autrement dit il est possible d'aller en n étapes de l'état i à l'état j si et seulement si on peut trouver dans le graphe G un chemin de longueur n allant de i à j.

On en déduit que

$$\mathbb{P}^i(\exists n > 0;\ X_n = j) = \mathbb{P}^i(\bigcup_{n\geq 1} \{X_n = j\}),$$

qui représente la probabilité que, partant de i, on puisse arriver à j, est non nulle si et seulement si il existe dans le graphe G un chemin allant de i à j. Dans ce cas, on dit que j est *accessible* à partir de i et on écrit $i \to j$.

Si il y a à la fois un chemin de i vers j et un chemin de j vers i, on dit que les états i et j communiquent et on écrit $i \leftrightarrow j$.

Si tous les états communiquent, on dit que la chaîne de Markov est *irréductible*.

On appelle *période* d'un état x d'une chaîne de Markov et on note $d(x)$ le pgcd (plus grand commun diviseur) des longueurs des circuits du graphe G contenant x. Lorsque la période est 1, on dit que l'état x est *apériodique*.

Lemme 12. *Si deux états communiquent, alors ils ont même période.*

Démonstration. Soient i, j avec $i \leftrightarrow j$. Soit γ un chemin de i à j, γ' un chemin de j à i. Soit $\mathcal{C}$ un circuit quelconque (éventuellement vide) contenant j . $\gamma - \gamma'$ et $\gamma - \mathcal{C} - \gamma'$ sont deux circuits contenant i. Donc $d(i)$ divise leurs longueurs ainsi que la différence de leurs longueurs, soit la longueur de $\mathcal{C}$. Ainsi $d(i)$ divise les longueurs de tous les circuits contenant j, donc divise leur pgcd, soit $d(j)$. De la même manière, on montre que $d(j)$ divise $d(i)$, d'où $d(i) = d(j)$. ✓

Définition: Si une chaîne irréductible a ses états de période 1, on dit qu'elle est apériodique.

Le lemme suivant et ses corollaires se révéleront très utiles par la suite.

Lemme 13. *Soit x un état de période 1. Il existe un entier $N(x)$ tel que pour tout $n \geq N(x)$ le graphe associé à la chaîne de Markov possède un circuit de longueur n contenant x*

Soit A l'ensemble des valeurs de n telles que le graphe associé à la chaîne de Markov possède un circuit de longueur n contenant x. Il est clair que A est stable par addition (concaténation des circuits). Il existe $p \geq 1$ et $n_1, n_2, \ldots, n_p$ tels que le pgcd de $n_1, n_2, \ldots, n_p$ soit 1. D'après le lemme de Bezout, il existe des relatifs $a_1, \ldots a_p$ tels que $1 = \sum_{k=1}^{p} a_k n_k$. Posons $P = \sum_{p: a_p > 0} a_p n_p$ et $N = \sum_{p: a_p < 0} (-a_p) n_p$. On a $P \in A, N \in A$ et $1 = P - N$. Soit $n \geq N(N-1)$. On peut écrire $n = bN + r$, avec $b \geq N - 1$ et $r \in \{0, 1, \ldots N-1\}$. On a $n = bN + r = bN + r(P - N) = rP + (b - r)N \in A$ car $b - r \in \mathbb{N}$ et A est stable par addition.

Corollaire 19. *Soit x est un état de période 1. On suppose qu'il existe un chemin de longueur $d(x, y)$ allant de x à y. Alors pour tout entier n tel que*

$$n \geq N(x, y) = N(x) + d(x, y),$$

il existe un chemin de longueur n allant de x à y. Ainsi, si P est la matrice associée, $P^n(x, y) > 0$.

Démonstration. Il suffit de concaténer le chemin allant de x à x avec un chemin allant de x à y. ✓

Corollaire 20. *Si une chaîne de Markov est irréductible, apériodique, à valeurs dans un ensemble fini S, alors il existe un entier N tel que pour tout $n \ge N$ et tout couple (i,j), il existe un chemin de longueur n allant de i à j. Ainsi, si P est la matrice associée, P^n est à coefficients strictement positifs.*

Démonstration. Il suffit de prendre $N = \max(N(x), x \in S) + \mathrm{diam}(G)$. ✓

La définition suivante est très simple, mais sera abondamment utilisée dans les exercices.

Définition: On appelle point absorbant (ou état absorbant) d'une chaîne tout point x tel que $\mathbb{P}^x(X_1 = x) = 1$.

6.3.5 Point de vue fonctionnel (*)

Une matrice stochastique indexée par S peut assez naturellement être vue comme un opérateur sur l'espace des fonctions bornées sur S. Rappelons que l'espace des fonctions bornées, que l'on note $\ell^\infty(S)$, est l'ensemble des fonctions f telles que

$$\|f\|_\infty = \sup\{|f(i)|; i \in S\} < +\infty,$$

et que $\|\cdot\|_\infty$ est une norme sur $\ell^\infty(S)$.

Maintenant, si $f \in \ell^\infty(S)$, la fonction Pf définie par

$$\forall i \in S \quad (Pf)(i) = \sum_{j \in S} p_{i,j} f(j)$$

est une fonction bornée. En effet, la majoration

$$\sum_{j \in S} |p_{i,j} f(j)| \le \sum_{j \in S} p_{i,j} \|f\|_\infty = \|f\|_\infty$$

montre que la série converge absolument. De plus, comme pour tout $i \in S$, $|Pf(i)| \le \|f\|_\infty$, on a $\|Pf\|_\infty \le \|f\|_\infty$: P est une contraction de l'espace des fonctions bornées.

Théorème 61. *Soit $(X_n)_{n \ge 0}$ une suite de variables aléatoires à valeurs dans S. On a équivalence entre :*

1. *$(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de passage P à valeurs dans S.*
2. *Pour toute fonction $f \in \ell^\infty(S)$ et pour tout entier $n \geq 0$*

$$\mathbb{E}[f(X_{n+1})|X_0,\dots,X_n] = (Pf)(X_n).$$

Démonstration. Pour le sens direct, il suffit de prendre $f = \delta_i$ et d'appliquer la caractérisation vue en début de chapitre. Regardons la réciproque. Si $f = \delta_i$, l'identité découle encore de la caractérisation vue en début de chapitre. Par linéarité, l'identité s'étend au cas des fonctions à support fini. Passons au cas dénombrable. Il existe une suite croissante d'ensembles finis $(S_p)_{p\geq 1}$ avec $S = \cup_{p\geq 1} S_p$. $f\mathbb{1}_{S_p}$ est une fonction bornée, donc

$$\mathbb{E}[(f\mathbb{1}_{S_p})(X_{n+1})|X_0,\dots,X_n] = (P(f\mathbb{1}_{S_p})(X_n).$$

Le théorème de convergence dominée pour l'espérance conditionnelle nous dit que

$$\lim_{p\to+\infty} \mathbb{E}[(f\mathbb{1}_{S_p})(X_{n+1})|X_0,\dots,X_n] = \mathbb{E}[f(X_{n+1})|X_0,\dots,X_n].$$

Pour conclure, il suffit de montrer que pour tout $i \in S$

$$\lim_{p\to+\infty} P(f\mathbb{1}_{S_p})(i) = P(f)(i).$$

Mais pour toute fonction bornée, le théorème de transfert donne

$$(Pg)(i) = \mathbb{E}^i g(X_1). \tag{6.4}$$

L'identité voulue découle alors immédiatement du théorème de convergence dominée. ✓

Remarque: La quantité $(Pf)(i)$ est l'intégrale de la fonction f par rapport à la mesure de probabilité sur S qui affecte à j la probabilité $p_{i,j}$. Cette propriété pourra, à l'occasion, être utile. Il en est de même de l'identité (6.4).

On déduit du théorème 61 une autre remarque très simple, mais très puissante :

Corollaire 21. *Soit $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de passage P à valeurs dans S, $f \in \ell^\infty(S)$. Si l'on pose $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$, alors la suite $(Y_n)_{n\geq 1}$ définie par*

$$\forall n \geq 0 \quad Y_{n+1} = f(X_{n+1}) - (Pf)(X_n)$$

est une suite de différences de martingales adaptée à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

Démonstration. Y_{n+1} est $\mathcal{F}_{n+1}$-mesurable et

$$\begin{aligned}\mathbb{E}[Y_{n+1}|\mathcal{F}_n] &= \mathbb{E}[f(X_{n+1})|\mathcal{F}_n] - \mathbb{E}[(Pf)(X_n)|\mathcal{F}_n] \\ &= (Pf)(X_n) - (Pf)(X_n) = 0.\end{aligned}$$

✓

Cette différence de martingale sera utilisée plusieurs fois dans le cours pour démontrer des résultats un peu avancés (voir notamment le théorème 85, ainsi que le théorème central limite proposé en problème corrigé).

Pour en voir un peu plus sur l'aspect fonctionnel, on pourra regarder le problème déjà cité ainsi que l'exercice 94.

6.4 Propriété de Markov

La propriété de Markov, que nous allons étudier maintenant, est, comme son nom l'indique, une propriété importante des chaînes de Markov. Elle exprime le fait que le comportement aléatoire d'une chaîne de Markov à partir d'un certain instant n'est pas influencé par le passé autrement que par l'état présent de la chaîne, que la valeur des états passés à contribué à déterminer. Cette propriété est très naturelle si l'on pense à la représentation d'une chaîne de Markov comme système dynamique aléatoire que l'on a rencontrée au corollaire 16. De fait, la preuve que l'on en propose ici va s'appuyer sur cette représentation.

6.4.1 Le théorème

Théorème 62. *Soit $(X_k)_{k\geq 0}$ une chaîne de Markov de matrice de passage P. Soit p un entier naturel. La suite $(X_{k+p})_{k\geq 0}$ est une chaîne de Markov de matrice de passage P et de loi initiale la loi de X_p. De plus, pour tout A $\mathcal{F}_p$-mesurable et tout $i\in S$, on a*

$$\mathbb{P}(A, X_p = i, X_{p+.}\in B) = \mathbb{P}(A, X_p = i)\mathbb{P}^i(X\in B).$$

De manière équivalente, on a $\mathbb{P}$-presque sûrement :

$$\mathbb{P}(X_{p+.}\in B|\mathcal{F}_p) = f_B(X_p), \text{ avec } f_B(x) = \mathbb{P}^x(X\in B).$$

Démonstration. Comme être une chaîne de Markov est une propriété de la loi, on peut supposer que $(X_n)_{n\geq 0}$ est obtenue par le procédé décrit plus haut : $X_{n+1} = f_{n+1}(X_n)$ où $(f_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes de loi χ_M,$(f_n)_{n\geq 1}$ étant de plus supposée indépendante de X_0. Posons $Y_n = X_{n+p}$ Si l'on pose encore $g_n = f_{n+p}$, on a la récurrence $Y_{n+1} = g_{n+1}(Y_n)$. Mais la loi de $(g_n)_{\geq 1}$ est $\chi_M^{\otimes}\mathbb{N}^*$, ce qui montre bien que $(Y_n)_{\geq 0}$ est une chaîne de Markov de matrice de passage P et de loi initiale la loi de X_p. Maintenant, soit B un borélien de $S^{\mathbb{N}}$. On pose

$$G_i((h_n)_{n\geq 1}) = (i, h_1(i), h_2\circ h_1(i), h_3\circ h_2\circ h_1(i), \dots).$$

On a alors

$$\mathbb{P}(A, X_p = i, X_{p+.}\in B) = \mathbb{P}(A, X_p = i, G_i(g)\in B).$$

$A\cap\{X_p = i\}$ est $\sigma(X_0, f_1, \dots, f_p)$-mesurable tandis que $\{G_i(g)\in B\}$ est $\sigma(f_k; k > p)$-mesurable, donc

$$\begin{aligned}\mathbb{P}(A, X_p = i, G_i(g)\in B) &= \mathbb{P}(A, X_p = i)\mathbb{P}(G_i(g)\in B)\\ &= \mathbb{P}(A, X_p = i)\mathbb{P}^i(X\in B).\end{aligned}$$

Pour la deuxième forme, il est clair que $f_B(X_p)$ est $\mathcal{F}_p$-mesurable ; il suffit donc de vérifier donc que pour tout $A\in\mathcal{F}_p$, on a

$$\mathbb{E}(\mathbb{1}_A\mathbb{1}_{\{X_{p+.}\in B\}}) = \mathbb{E}(\mathbb{1}_A f_B(X_p)).$$

Comme tout est positif, le théorème de Tonelli s'applique et on a

$$\mathbb{E}(\mathbb{1}_A f_B(X_p)) = \sum_{i \in S} \mathbb{E}(\mathbb{1}_A \mathbb{1}_{\{X_p=i\}} \mathbb{1}_{\{X_{p+.} \in B\}}).$$

Par ailleurs, pour tout $i \in S$, on a

$$\begin{aligned}\mathbb{E}(\mathbb{1}_A \mathbb{1}_{\{X_p=i\}} \mathbb{1}_{\{X_{p+.} \in B\}}) &= \mathbb{P}(A, X_p = i, X_{p+.} \in B)\\ &= \mathbb{P}(A, X_p = i)\mathbb{P}^i(X \in B)\\ &= \mathbb{E}[\mathbb{1}_{A \cap \{X_p=i\}} \mathbb{P}^i(X \in B)]\\ &= \mathbb{E}[\mathbb{1}_{A \cap \{X_p=i\}} f_B(X_p)],\end{aligned}$$

ce qui donne en resommant

$$\mathbb{E}(\mathbb{1}_A \mathbb{1}_{\{X_{p+.} \in B\}}) = \mathbb{E}[\mathbb{1}_A f_B(X_p)].$$

✓

Remarque: on peut trouver dans la littérature l'écriture

$$\mathbb{P}(X_{p+.} \in B | \mathcal{F}_p) = \mathbb{P}^{X_p}(X \in B).$$

Je mets le lecteur en garde contre le fait que $\mathbb{P}^{X_p}$ ne signifie pas la même chose que $\mathbb{P}^{\mathbb{P}_{X_p}}$.

La propriété de Markov est souvent utilisée sous la forme simple suivant : si A est un borélien de $\mathbb{R}^n$, B un borélien de $\mathbb{R}^p$, alors

$$\begin{aligned}&\mathbb{P}((X_0, \dots X_{n-1}) \in A, X_n = i, (X_{n+1}, \dots, X_{n+p}) \in B)\\ &= \mathbb{P}((X_0, \dots X_{n-1}) \in A, X_n = i)\mathbb{P}^i((X_1, \dots, X_p \in B)).\end{aligned}$$

6.4.2 Analyse au premier pas

On va maintenant décrire quelques variations autour d'une technique, dite de l'analyse au premier pas. Il s'agit essentiellement de découper l'événement que l'on veut analyser suivant la position de la chaîne au temps un, et d'appliquer la propriété de Markov. Malgré sa simplicité, la technique est redoutablement efficace, particulièrement dans l'étude d'événements de la tribu de queue de la chaîne.

Corollaire 22. *Soit $(X_n)_{n\geq 0}$ une chaîne de Markov de matrice de passage $(p_{i,j})$, B un borélien de $\mathbb{R}^{\mathbb{N}}$. On note θ l'opérateur de translation : $\theta((x_n)_{n\geq 0}) = ((x_{n+1})_{n\geq 0})$.*

Alors

$$\begin{aligned}\mathbb{P}(\theta(X) \in B) &= \mathbb{P}^{\mathbb{P}_{X_1}}(X \in B) \\ &= \sum_{j:\mathbb{P}(X_1=j)>0} \mathbb{P}(X_1 = j)\mathbb{P}^j(X \in B).\end{aligned}$$

En particulier, si B est invariant par l'opérateur de translation (c'est à dire que $\theta^{-1}(B) = B$) , alors on a le système d'équations :

$$\mathbb{P}^i(X \in B) = \sum_{j:p_{i,j}>0} p_{i,j}\mathbb{P}^j(X \in B).$$

Démonstration. La première égalité traduit exactement la propriété de Markov : une chaîne de Markov observée à partir du temps 1 a la même loi qu'une chaîne de Markov de même dynamique commençant avec comme valeur initiale celle que prend la chaîne de Markov non décalée au temps 1. La deuxième égalité correspond à une décomposition suivant les valeurs que peut prendre X_1.

Passons au cas où B est invariant :

$$\begin{aligned}\mathbb{P}^i(X \in B) &= \mathbb{P}^i(X \in \theta^{-1}(B)) \\ &= \mathbb{P}^i(\theta(X) \in B) \\ &= \sum_{i:P^i(X_1=j)>0} \mathbb{P}^i(X_1 = j)\mathbb{P}^j(X \in B) \\ &= \sum_{i:p_{i,j}>0} p_{i,j}\mathbb{P}^j(X \in B).\end{aligned}$$

✓

Le corollaire précédent prend une forme particulièrement intéressante lorsque la chaîne de Markov est réalisée sur l'espace canonique.

Corollaire 23. *Une dynamique markovienne sur D dénombrable étant fixée, on note $\mathbb{P}^i$ la mesure markovienne associée sur $D^{\mathbb{N}}$ partant de l'état*

$i \in D$. Soit $\mathbb{P}$ une mesure markovienne quelconque associée à la dynamique. Alors, si θ désigne l'opérateur de translation sur $D^{\mathbb{N}}$, on a

$$\mathbb{P}_\theta = \sum_{i \in D} \mathbb{P}(X_1 = i)\mathbb{P}^i.$$

En particulier, si $(X_n)_{n \geq 0}$ est la suite des projections canoniques, on a pour tout $i \in D$, $\mathbb{P}(\cdot|X_0 = i) = \mathbb{P}^i$.

Démonstration. La première identité est un cas particulier du théorème précédent. En appliquant cette égalité à l'événement $B \cap \{X_0 = k\}$, on a $\mathbb{P}_\theta(B, X_0 = k) = \mathbb{P}(X_1 = k)\mathbb{P}^k(B, X_0 = k)$ (les autres termes de la somme sont nuls). Comme $X_1 = X_0 \circ \theta$, $\mathbb{P}_\theta(X_0 = k) = \mathbb{P}(X_1 = k)$, d'où le résultat voulu.

✓

Remarque : on peut également démontrer ce dernier corollaire directement, en procédant comme dans l'exercice 50.

Corollaire 24. *Une dynamique markovienne sur D dénombrable étant fixée, on note $\mathbb{P}^i$ la mesure markovienne associée sur $D^{\mathbb{N}}$ partant de $i \in D$. Soit $\mathbb{P}$ une mesure markovienne quelconque associée à la dynamique. Alors, si θ désigne l'opérateur de translation sur $D^{\mathbb{N}}$, on a pour toute fonction F mesurable de $D^{\mathbb{N}}$ dans $[0, +\infty]$ et pour tout $i \in D$:*

$$\mathbb{E}(\mathbb{1}_{\{X_1=i\}}(F \circ \theta)) = \mathbb{P}(X_1 = i)\mathbb{E}^i(F).$$

Démonstration. Posons $G = \mathbb{1}_{\{X_0=i\}} \times F$. On a $\mathbb{1}_{\{X_1=i\}}(F \circ \theta) = G \circ \theta$, donc

$$\begin{aligned}\mathbb{E}(\mathbb{1}_{\{X_1=i\}} F \circ \theta)) &= \mathbb{E}(G \circ \theta) \\ &= \int G \, d\mathbb{P}_\theta \\ &= \sum_{j \in D} \mathbb{P}(X_1 = j)\mathbb{E}^j(G),\end{aligned}$$

où la dernière égalité vient du corollaire 23 et de la formule d'intégration par rapport à une mesure qui s'écrit comme une série de mesures.

Pour $j = i$, on a $G = F$ $\mathbb{P}^j$-presque partout, donc $\mathbb{E}^j(G) = \mathbb{E}^i(F)$, et pour $j \neq i$, $G = 0$ $\mathbb{P}^j$-presque partout, donc $\mathbb{E}^j(G) = 0$. La formule annoncée en découle immédiatement. ✓

6.5 L'analyse au premier pas en action (*)

Cette section peut être omise en première lecture.

Dans cette section, on étudie le temps d'entrée dans un ensemble A d'une chaîne de Markov $(X_n)_{n\geq 0}$ à valeurs dans D fini. Afin de simplifier les écritures, on suppose que $(X_n)_{n\geq 0}$ est défini sur l'espace canonique ; on dispose ainsi de l'opérateur de translation θ.

On rappelle que si A est une partie de S, le temps d'entrée dans A, noté T_A, est défini par

$$T_A = \inf\{n \geq 0; X_n \in A\}.$$

A priori, T_A prend ses valeurs dans $\mathbb{N} \cup \{+\infty\}$. Notons que si le point de départ de la chaîne n'est pas dans A, T_A coïncide avec

$$\begin{aligned}
&\inf\{n \geq 1; X_n \in A\} \\
&= 1 + \inf\{n \geq 0; X_{n+1} \in A\} \\
&= 1 + T_A \circ \theta.
\end{aligned}$$

6.5.1 Analyse au premier pas et fonction de répartition

Pour $i \in S\backslash A$, on a $\mathbb{P}^i$ presque sûrement $T_A = T_A \circ \theta + 1$. On en déduit

$$\begin{aligned}
\mathbb{P}^i(T_A \leq n+1) &= \mathbb{P}^i(T_A \circ \theta \leq n) \\
&= \mathbb{P}^i_\theta(T_A \leq n) \\
&= \sum_{j \in S} \mathbb{P}^i(X_1 = j)\mathbb{P}^j(T_A \leq n) \\
&= \sum_{j \in S\backslash A} p_{i,j}\mathbb{P}^j(T_A \leq n) + \sum_{j \in A} p_{i,j}. \qquad (6.5)
\end{aligned}$$

Si l'on note $(p_{i,j}) = \begin{pmatrix} M & N \\ * & * \end{pmatrix}$ en notant à la fin les éléments de A et $F_n = (\mathbb{P}^i(T_A \leq n))_{i \in S \setminus A}$, on a alors

$$F_{n+1} = MF_n + N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|}, \tag{6.6}$$

De même,

$$F_{n+2} = MF_{n+1} + N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|},$$

et en faisant la différence, $(F_{n+1} - F_{n+2}) = M(F_n - F_{n+1})$, soit

$$(\mathbb{P}(T_A = n+1))_{i \in S \setminus A}) = M(\mathbb{P}(T_A = n))_{i \in S \setminus A}),$$

d'où l'on déduit aisément que pour tout $n \geq 1$,

$$(\mathbb{P}(T_A = n))_{i \in S \setminus A}) = M^{n-1} N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|} \tag{6.7}$$

6.5.2 Analyse au premier pas et probabilité des issues possibles

Soit $i \in S \setminus A$. En faisant tendre n vers l'infini, l'équation (6.5) et le théorème de continuité séquentielle croissante donnent

On a

$$\mathbb{P}^i(T_A < +\infty) = \sum_{j \in S \setminus A} p_{i,j} \mathbb{P}^j(T_A < +\infty) + \sum_{j \in A} p_{i,j},$$

Posant $F = (\mathbb{P}^i(T_A < +\infty))_{i \in S \setminus A}$, on a alors

$$F = MF + N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|}, \tag{6.8}$$

Comme la matrice est stochastique, on a

$$M \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|S\backslash A|} + N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|S\backslash A|} ,$$

On peut noter l'identité

$$N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|} = (I - M) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|S\backslash A|} , \tag{6.9}$$

qui nous sera encore utile. On obtient ainsi

$$(I - M)F = (I - M) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|S\backslash A|} . \tag{6.10}$$

Si $I - M$ est inversible, on obtient ainsi $\mathbb{P}^i(T_A < +\infty) = 1$ pour tout $i \in S\backslash A$.[3] Sinon, des équations supplémentaires sont requises pour déterminer le système. C'est typiquement ce qui arrive lorsque $S\backslash A$ contient un ensemble absorbant, disons B. Dans ce cas, les équations $F_i = 0$ pour $i \in B$ peuvent parfois suffire à déterminer le système.

Supposons par exemple que les matrices M et N s'écrivent

$$M = \begin{pmatrix} M_1 & M_2 \\ 0 & S \end{pmatrix} \text{ et } N = \begin{pmatrix} N_1 \\ 0 \end{pmatrix},$$

où S est une matrice stochastique correspondant aux entrées de l'ensemble B. Alors, B est absorbant et une chaîne de Markov partant d'un point de B reste toujours dans B, ce qui entraîne que pour $i \in B$ $\mathbb{P}^i(T_A < +\infty) = 0$: si l'on pose $F_1 = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$, avec F_2 correspondant

3. Il faut être conscient que tester l'inversibilité de $I - M$ n'est pas une méthode efficace pour déterminer si $\mathbb{P}^i(T_A < +\infty) = 1$. Ce type de problème se résout par la classification des états de la chaîne, qui sera abordée dans le chapitre suivant.

aux entrées de B, on a $F_2 = 0$. Ainsi, comme

$$MF + N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|} = \begin{pmatrix} M_1 F_1 + N_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|} \\ 0 \end{pmatrix},$$

$$F_1 = M_1 F_1 + N_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|}, \tag{6.11}$$

Ainsi, si $I - F_1$ est inversible, on a

$$F_1 = (I - M_1)^{-1} N_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|}.$$

On peut toutefois faire une remarque théorique, valable en toute généralité : Si f^0 est une solution à coefficients positifs de l'équation (6.8), on voit facilement que l'ensemble

$$U_{f^0} = \{x \in \mathbb{R}_+^{S \backslash A}; \forall i \in S \backslash A x_i \leq f_i^0\}$$

est stable par l'application $x \mapsto Mx + N \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{|A|}$.

Comme $F_0 = 0 \in U_{f_0}$, avec (6.6), on obtient $F_n \in U_{f_0}$ pour tout n, d'où à la limite $F \in U_{f_0}$. Il s'ensuit que F est la plus petite solution de l'équation (6.8) à coefficients positifs.

6.5.3 Analyse au premier pas et temps moyen d'atteinte

Pour $i \in S\backslash A$, on a $\mathbb{P}^i$ presque sûrement

$$\begin{aligned}\mathbb{1}_{\{T_A<+\infty\}}T_A &= \mathbb{1}_{\{T_A<+\infty\}}(1+T_A\circ\theta)\\ &= \mathbb{1}_{\{T_A<+\infty\}}+\mathbb{1}_{\{T_A<+\infty\}}(T_A\circ\theta)\\ &= \mathbb{1}_{\{T_A<+\infty\}}+\mathbb{1}_{\{T_A\circ\theta<+\infty\}}(T_A\circ\theta)\\ &= \mathbb{1}_{\{T_A<+\infty\}}+(\mathbb{1}_{\{T_A<+\infty\}}T_A)\circ\theta,\end{aligned}$$

d'où

$$\begin{aligned}\mathbb{E}^i(\mathbb{1}_{\{T_A<+\infty\}}T_A) &= \mathbb{E}^i(\mathbb{1}_{\{T_A<+\infty\}})+\mathbb{E}_{\mathbb{P}^i_\theta}(\mathbb{1}_{\{T_A<+\infty\}}T_A)\\ &= \mathbb{P}^i(T_A<+\infty)+\sum_{j\in S}p_{i,j}\mathbb{E}^j(\mathbb{1}_{\{T_A<+\infty\}}T_A)\\ &= \mathbb{P}^i(T_A<+\infty)+\sum_{j\in S\backslash A}p_{i,j}\mathbb{E}^j(\mathbb{1}_{\{T_A<+\infty\}}T_A).\end{aligned}$$

Si l'on note $E=(\mathbb{E}^j(\mathbb{1}_{\{T_A<+\infty\}}T_A))_{j\in S\backslash A}$, on a alors

$$E=F+ME, \tag{6.12}$$

soit si $I-M$ inversible

$$E=(I-M)^{-1}F=(I-M)^{-2}N\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}_{|A|}. \tag{6.13}$$

En procédant comme précédemment, l'identité

$$\mathbb{E}^i(\mathbb{1}_{\{T_A\le n+1\}}T_A)=\mathbb{P}^i(T_A<+\infty)+\sum_{j\in S\backslash A}p_{i,j}\mathbb{E}^j(\mathbb{1}_{\{T_A\le n\}}T_A)$$

permet de montrer qu'en toute généralité, E est la plus petite solution à coefficients positifs de l'équation $e=F+Me$.

6.5.4 Passage aux fonctions génératrices

Rappelons que la fonction génératrice d'une variable aléatoire entière X est définie sur $[-1,1]$ par

$$\varphi_X(t) = \mathbb{E}(t^X) = \sum_{k=0}^{+\infty} \mathbb{P}(X=k)t^k.$$

Il s'agit d'une série entière, de rayon de convergence supérieur ou égal à 1 donc les valeurs de fonction permettent de retrouver les coefficients puisque l'on a

$$\frac{\varphi_X^{(n)}(0)}{n!} = \mathbb{P}(X=n).$$

On a toujours $\varphi_X(1)=1$. Si le rayon de convergence est supérieur à un, il est facile de voir qu'on a l'expression des moments factoriels [4] :

$$\varphi_X^{(n)}(1) = \mathbb{E}X(X-1)\dots(X-n+1).$$

Voyons comment appliquer ça aux chaînes de Markov.

Maintenant, pour tout $i \in S$, on pose $\varphi_i(x) = \sum_{n=0}^{+\infty} \mathbb{P}^i(T_A=n)x^n$. φ_i est la fonction génératrice de $T_A 1\!\!1_{T_A<+\infty}$, et permet ainsi de retrouver la loi de $T_A 1\!\!1_{T_A<+\infty}$ sous $\mathbb{P}^i$.

En particulier $\varphi^i(1) = \mathbb{P}^i(T_A<+\infty)$ et $(\varphi^i)'(1) = \mathbb{E}^i(T_A 1\!\!1_{\{T_A<+\infty\}})$.

Calculons-la encore par une nouvelle analyse au premier pas.

4. De manière générale, il est encore vrai que X admet un moment d'ordre n si et seulement si φ_X est n fois dérivable à gauche en 1, et l'identité des moments factoriels est encore vérifiée. Pour une preuve, on peut s'inspirer des calculs qui seront développés page 205.

Pour $i \in A$, $\varphi_i(x) = 1$, tandis que pour $i \in S\backslash A$, on a

$$\begin{aligned}
\varphi_i(x) &= \mathbb{E}^i(x^{T_A(x)}) \\
&= \mathbb{P}^i(T_A = 0) + \sum_{n=0}^{+\infty} \mathbb{P}^i(T_A = n+1)x^{n+1} \\
&= x\sum_{n=0}^{+\infty}\sum_{j\in S} \mathbb{P}^i(X_1 = j, T_A\circ\theta = n)x^n \\
&= x\sum_{n=0}^{+\infty}\sum_{j\in S} p_{i,j}\mathbb{P}^j(T_A = n)x^n \\
&= x\sum_{j\in S} p_{i,j}\varphi^j(x) \\
&= x\sum_{j\in S\backslash A} p_{i,j}\varphi^j(x) + x\sum_{j\in A} p_{i,j}.
\end{aligned}$$

Si l'on note $\varphi = (\varphi_i)_{i\in S\backslash A}$, on a alors

$$\varphi(x) = xM\varphi(x) + xN\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix},$$

En évaluant φ au moint $x = 1$, on retrouve la formule (6.8) Si $I - xM$ est inversible, on a alors

$$\varphi(x) = x(1 - xM)^{-1}N\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}_{|A|}.$$

Posons, pour x tel que $I - xM$ est inversible, $\psi(x) = (1 - xM)^{-1}$. Alors, pour h suffisamment petit, on a

$$\begin{aligned}
\psi(x+h) &= (1 - xM)^{-1} = ((I - xM) - hM)^{-1} \\
&= (1 - xM)^{-1}(I - h(I - xM)^{-1}M)^{-1} \\
&= (1 - xM)^{-1}(I + h(I - xM)^{-1}M + o(h)).
\end{aligned}$$

On a donc $\psi'(x) = \psi(x)^2 M$; cela permettra de calculer les dérivées successives. On peut noter que

$$\begin{aligned}(x\psi)'(x) &= x\psi'(x) + \psi(x) \\ &= \psi(x)^2 Mx + \psi(x) \\ &= \psi(x)^2 + \psi(x)^2(Mx - I) + \psi(x) \\ &= \psi(x)^2,\end{aligned}$$

d'où $\varphi'(x) = \psi(x)^2 N \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}_{|A|}$. Par récurrence, il vient aisément

$$\varphi^{(n)}(x) = n!\psi(x)^{n+1} M^{n-1} N \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}_{|A|}.$$

Supposons maintenant que $I - M$ est inversible : en évaluant au point 1, on obtient les moments factoriels :
Si $E_n = (\mathbb{E}^i T_A(T_A - 1)\dots(T_A - n + 1)\mathbb{1}_{T_A < +\infty})_{i \in S\setminus A}$, alors

$$E_n = \varphi^{(n)}(1) = n!(I - M)^{-(n+1)} M^{n-1} N \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}_{|A|}.$$

En particulier

$$E = E_1 = \varphi'(1) = (1 - M)^{-2} N \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}_{|A|},$$

et on retrouve ainsi la formule (6.13).

On a enfin accès aux probabilités des différentes valeurs pour le temps d'atteinte :

$$(\mathbb{P}^i(T_A = n))_{i \in S\setminus A} = \left(\frac{\varphi_i^{(n)}(0)}{n!}\right)_{i \in S\setminus A} = M^{n-1} N \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}_{|A|}$$

et on retrouve ainsi la formule (6.7).

6.5.5 Quelques exemples

Temps d'absorption et probabilités des issues : calculs avec Sage

A et B se confrontent dans un jeu, la probabilité que A gagne une manche est a, que B gagne une manche est b et qu'ils fassent manche nulle est $c = 1 - a - b$. Est déclaré gagnant celui qui gagne 2 manches consécutives pour la première fois. Calculer la probabilité que A gagne.

Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes avec $\mathbb{P}(Y_1 = 2) = a, \mathbb{P}(Y_1 = 3) = b, \mathbb{P}(Y_1 = 1) = c = 1 - a - b$. On note

$$H_n^1 = \{\text{la partie est en cours temps } n\text{, manche } n \text{ sans gagnant}\}$$
$$H_n^2 = \{\text{la partie est en cours au temps } n\text{, } A \text{ a gagné la manche } n\}$$
$$H_n^3 = \{\text{la partie est en cours au temps } n\text{, } B \text{ a gagné la manche } n\}$$
$$H_n^4 = \{\text{la partie est terminée au temps } n \text{ et } B \text{ a gagné}\}$$
$$H_n^5 = \{\text{la partie est terminée au temps } n \text{ et } A \text{ a gagné}\}$$

Si je pose $X_n = \sum_{k=1}^5 k\mathbb{1}_{\{H_n^k\}}$, on a la récurrence $X_{n+1} = F(X_n, Y_{n+1})$, avec $F(x,y) = x$ pour $x \in \{4,5\}$, $F(1,y) = y$, $F(x,x) = 7 - x$ pour $x \in \{2,3\}$, $F(x,y) = y$ pour $x \in \{2,3\}$ et $x \neq y$. La chaîne de Markov admet la matrice de passage

$$\begin{pmatrix} 1-a-b & a & b & 0 & 0 \\ 1-a-b & 0 & b & 0 & a \\ 1-a-b & a & 0 & b & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

La probabilité que A gagne est $\mathbb{P}^1(T_A < +\infty)$, avec $A = \{5\}$. Si on pose $B = \{4\}$, B est un ensemble absorbant, on a ici

$$M_1 = \begin{pmatrix} 1-a-b & a & b \\ 1-a-b & 0 & b \\ 1-a-b & a & 0 \end{pmatrix} \text{ et } N_1 = \begin{pmatrix} 0 \\ a \\ 0 \end{pmatrix}.$$

Ainsi, en admettant que $I - M_1$ est inversible, on a

$$\begin{pmatrix} \mathbb{P}^1(A \text{ gagne }) \\ \mathbb{P}^2(A \text{ gagne }) \\ \mathbb{P}^3(A \text{ gagne }) \end{pmatrix} = (I - M_1)^{-1} N_1.$$

Confions le calcul à Sage.

```
sage: a=var('a')
sage: b=var('b')
sage: M1=matrix([[1-a-b,a,b],[1-a-b,0,b],[1-a-b,a,0]])
sage: factor(det(1-M1))
a^2*b + a*b^2 + a^2 + b^2
sage: # donc la matrice est inversible si a>0 ou b>0
sage: N1=matrix([[0],[a],[0]])
sage: P=(1-M1)^(-1)*N1
sage: factor(P)
[      a^2*(b + 1)/(a^2*b + a*b^2 + a^2 + b^2)]
[(a*b + b^2 + a)*a/(a^2*b + a*b^2 + a^2 + b^2)]
[              a^2/(a^2*b + a*b^2 + a^2 + b^2)]
```

Ce qui nous donne la probabilité de gain cherchée :

$$\frac{a^2(b+1)}{a^2b + ab^2 + a^2 + b^2}$$

Passons au calcul de la durée de la moyenne de la partie. On prend cette fois $A = \{4, 5\}$ et c'est la matrice M_1 qui joue le rôle de M. Comme $I - M_1$ est inversible, on a $F = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} \mathbb{E}^1(T) \\ \mathbb{E}^2(T) \\ \mathbb{E}^3(T) \end{pmatrix} = (I - M_1)^{-1} F$.

Confions le calcul à Sage.

```
sage: E=(1-M1)^(-1)*matrix([[1],[1],[1]])
sage: factor(E)
[(a + 1)*(b + 1)/(a^2*b + a*b^2 + a^2 + b^2)]
[        (b + 1)/(a^2*b + a*b^2 + a^2 + b^2)]
[        (a + 1)/(a^2*b + a*b^2 + a^2 + b^2)]
```

D'où le temps moyen cherché :

$$\frac{(a+1)(b+1)}{a^2b+ab^2+a^2+b^2}$$

On va enfin calculer la fonction caractéristique de T dans deux cas particuliers :

– Lorsque $a=b=p$:

```
sage: x=var('x')
sage: p=var('p')
sage: a=p
sage: b=p
sage: M1=matrix([[1-a-b,a,b],[1-a-b,0,b],[1-a-b,a,0]])
sage: r=x*(1-M1*x)^(-1)*(1-M1)*matrix([[1],[1],[1]])
sage: factor(r)
[          2*p^2*x^2/(2*p^2*x^2 - p*x^2 + p*x - x + 1)]
[(2*p*x - x + 1)*p*x/(2*p^2*x^2 - p*x^2 + p*x - x + 1)]
[(2*p*x - x + 1)*p*x/(2*p^2*x^2 - p*x^2 + p*x - x + 1)]
```

D'où la fonction génératrice : $\varphi^1(x) = \frac{2p^2x^2}{(2p^2-p)x^2+(p-1)x+1}$.

– Lorsque $a=p$ et $b=0$. Cela correspond au cas où on attend la première fois où on a deux réussites consécutives dans une suite d'expériences aléatoires indépendantes de même probabilité.

```
sage: x=var('x')
sage: p=var('p')
sage: a=p
sage: b=0
sage: M1=matrix([[1-a-b,a,b],[1-a-b,0,b],[1-a-b,a,0]])
sage: r=x*(1-M1*x)^(-1)*(1-M1)*matrix([[1],[1],[1]])
sage: factor(r)
[          p^2*x^2/(p^2*x^2 - p*x^2 + p*x - x + 1)]
[(p*x - x + 1)*p*x/(p^2*x^2 - p*x^2 + p*x - x + 1)]
[          p^2*x^2/(p^2*x^2 - p*x^2 + p*x - x + 1)]
```

D'où la fonction génératrice : $\varphi^1(x) = \frac{p^2x^2}{1-(1-p)x-p(1-p)x^2}$.

6.5.6 Un calcul de fonction génératrice

Voyons comment on peut calculer la fonction génératrice. On sait que

$$\begin{aligned}\varphi(x) &= x(1-xM)^{-1}N\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}_{|A|}\\ &= \frac{x}{\det(1-xM)}{}^t\mathrm{Com}(1-xM)N\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}_{|A|},\end{aligned}$$

Ainsi, il existe un polynôme Q_i de degré inférieur ou égal à $|S\backslash A|$, nul en 0 et tel que

$$\varphi_i(x) = \frac{Q_i(x)}{\det(1-xM)} = \frac{Q_i(x)}{\tilde{P}_M(x)},$$

où $\tilde{P}_M$ est le polynôme réciproque du polynôme caractéristique

$$P_M(X) = \det(X\mathrm{Id} - M).$$

Si on pose $n = |S\backslash A|$, on a

$$\det(1-xM) = x^n \det(1/x - M) = x^n P_M(1/x) = \tilde{P}_M(x).$$

Comme $\varphi_i(x) = \sum_{i=1}^{|S\backslash A|} \mathbb{P}(T_A = i)x^i + o(x^{|S\backslash A|})$, on a

$$Q_i(x) = \left(\sum_{i=1}^{|S\backslash A|} \mathbb{P}(T_A = i)x^i\right)\tilde{P}_M(x) + o(x^{|S\backslash A|}),$$

ce qui peut permettre parfois d'identifier les coefficients de Q_i.

Un cas particulièrement favorable pour le calcul de la fonction génératrice de T_A est celui où l'on part d'un site i tel que $\mathbb{P}^i(T_A < |S\backslash A|) = 0$ et $\mathbb{P}^i(T_A = |S\backslash A|) > 0$.

En 0, on a alors l'équivalent $\varphi_i(x) \sim Q_i(x)$.

Comme $\mathbb{P}^i(T_A < |S\backslash A|) = 0$ et $\mathbb{P}^i(T_A = |S\backslash A|) > 0$, on a aussi $\varphi_i(x) \sim \mathbb{P}(T_A = |S\backslash A|)x^{|S\backslash A|}$, d'où $Q_i(x) = \mathbb{P}(T_A = |S\backslash A|)x^{|S\backslash A|}$ vu le degré de Q_i, ce qui nous donne

$$\varphi_i(x) = \frac{x^n \mathbb{P}^i(T_A = n)}{\tilde{P}_M(x)}.$$

On va appliquer cette formule au calcul de la fonction génératrice de la première séquence de r réussites dans une suite d'expériences aléatoires indépendantes de même paramètre.

On pose $Y_0 = 0$, puis, pour $n \geq 1$, $(Y_n)_{n\geq 1}$ est une suite de variables indépendantes suivant la loi de Bernoulli de paramètre p. On pose ensuite $X_0 = 0$, et pour tout $n \geq 1$:

$$X_n = \max\{k : Y_n = Y_{n-1} = \dots Y_{n-k+1} = 1\} \wedge 0.$$

Si on pose $F(x, y) = (x + 1)y$, on a la relation $X_{n+1} = F(X_n, Y_{n+1})$, donc $(X_n)_{n\geq 1}$ est une chaîne de Markov. On arrête la chaîne dès qu'elle touche le point r.

La chaîne arrêtée, notée $(X'_n)_{n\geq 0}$, est encore une chaîne de Markov [5], avec la relation $X'_{n+1} = G(X'_n, Y_{n+1})$ où $G(x, y) = F(x, y)$ si $x \in \{0, 1, \dots, r-1\}$ et $G(r, y) = r$. Partant de 0, le premier temps où X_n touche r coïncide avec le premier temps où X'_n touche r. La fonction génératrice de

$$\inf\{n \geq 1; Y_n = Y_{n-1} = \dots = Y_{n-(r-1)} = 1\}$$

coïncide donc avec $\mathbb{E}^0(x^{T_r})$, où

$$T_r = \inf\{n \geq 1; X'_n = r\}.$$

5. Ceci est un fait général, voir l'exercice 62.

La chaîne (X'_n) a comme espace d'états $S = \{0, \dots, r\}$ et sa matrice est

$$\begin{pmatrix} 1-p & p & 0 & \dots & \dots & 0 & 0 \\ \vdots & 0 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 & \vdots \\ \vdots & \vdots & & & \ddots & p & 0 \\ 1-p & 0 & \dots & \dots & \dots & 0 & p \\ 0 & 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}.$$

En prenant $A = \{r\}$, avec les notations précédentes, on a

$$P_M(X) = \begin{pmatrix} X-(1-p) & -p & 0 & \dots & \dots & 0 \\ -(1-p) & X & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \dots & 0 & X & -p \\ -(1-p) & 0 & \dots & \dots & 0 & X \end{pmatrix}.$$

Pour calculer le déterminant, on fait l'opération

$$C_1 \leftarrow C_1 + \frac{p-1}{p-X}(C_2 + \dots C_r),$$

puis on développe suivant la première colonne : on obtient

$$\begin{aligned}
P_M(X) &= \left(X-(1-p)+\frac{p(1-p)}{p-X}\right)X^{r-1} \\
&\quad+\left(p-1+\frac{p-1}{p-X}X\right)(-p)^{r-1}(-1)^{r-1} \\
&= \left(X-(1-p)+\frac{p(1-p)}{p-X}\right)X^{r-1} \\
&\quad+\left(p-1+\frac{p-1}{p-X}X\right)p^{r-1} \\
&= \left(X-(1-p)\right)X^{r-1}+(p-1)p^{r-1}+\frac{1-p}{p-X}(pX^{r-1}-Xp^{r-1}) \\
&= \left(X-(1-p)\right)X^{r-1} \\
&\quad+\frac{1-p}{p-X}(pX^{r-1}-Xp^{r-1}+(X-p)p^{r-1}) \\
&= \left(X-(1-p)\right)X^{r-1}+\frac{p(1-p)}{p-X}(X^{r-1}-p^{r-1}) \\
&= \left(X-(1-p)\right)X^{r-1}-p(1-p)\frac{X^{r-1}-p^{r-1}}{X-p}.
\end{aligned}$$

D'où

$$\begin{aligned}
X^rP_M(1/X) &= (1-(1-p)X)-p(1-p)\frac{X^2-p^{r-1}X^{r+1}}{1-pX} \\
&= (1-(1-p)X)-p(1-p)X^2\frac{1-(pX)^{r-1}}{1-pX},
\end{aligned}$$

et finalement

$$\varphi(x)=\frac{p^rx^r}{1-(1-p)x-p(1-p)x^2\frac{1-(px)^{r-1}}{1-px}}.$$

Pour $r=2$, on retrouve le résultat de la section précédente.

6.6 Exercices sur les chaînes de Markov

6.6.1 Exercices corrigés

Exercice 62. *Chaîne de Markov arrêtée.*

Soit (X_n) une chaîne de Markov sur un ensemble dénombrable E, de matrice de transition Q. Étant donné un ensemble $B \subset E$, on note

$$T_B = \inf\{n \geq 0; X_n \in B\}$$

le temps d'entrée dans B et on pose $Y_n = X_{n \wedge T_B}$. Montrer que $(Y_n)_{n\geq 0}$ est une chaîne de Markov sur E dont on précisera la matrice de transition.

Exercice 63. *La ruine du joueur* ou *marche aléatoire avec barrière*.
Un joueur possédant une fortune de a unités joue à pile ou face jusqu'à ce qu'il ait fait sauter la banque ou qu'il soit ruiné. Les réserves de la banque sont de b unités. Chaque victoire rapporte une unité et chaque défaite en coûte une. On suppose que les lancers sont indépendants et que la probabilité de gain de la banque est $p = 1-q$. On veut déterminer la probabilité p_g que la banque résiste.

On note $(X_n)_{n\geq 1}$ une suite de v.a.i.i.d. de loi $p\delta_1 + q\delta_{-1}$, puis $S_n = \sum_{k=1}^n X_k$ et $T = \inf\{n \geq 0; S_n = -b \text{ ou } S_n = a\}$. Si l'on pose $S'_n = S_{n\wedge T}$, il est aisé de constater que S'_n représente la suite des gains relatifs de la banque.

1. Montrer que S'_n est une chaîne de Markov homogène à espace d'états $E = \{-b, \dots, a\}$ dont on déterminera la loi initiale et la matrice de transition.
2. Considérons les chaînes de Markov ayant la même matrice de transition que $(S'_n)_{n\geq 0}$ Montrer que la suite $(u_n)_{-b\leq n\leq a}$ définie par
$$u_n = \mathbb{P}^n(\{\text{la banque résiste}\})$$
vérifie la récurrence linéaire
$$pu_{n+1} - u_n + qu_{n-1} = 0.$$
Que valent u_a et u_{-b} ?

3. Résoudre l'équation de récurrence et en déduire

$$p_g = \begin{cases} \frac{(\frac{q}{p})^b - 1}{(\frac{q}{p})^{a+b} - 1} & \text{si } p \neq q \\ \frac{b}{a+b} & \text{si } p = q = \frac{1}{2}. \end{cases} \tag{6.14}$$

4. Dans cette question, $p = q = 1/2$. On note $v_n = \mathbb{E}^n[T]$. Montrer que si $-b < n < a$, on a

$$v_n = 1 + \frac{1}{2}(v_{n+1} + v_{n-1}),$$

puis exprimer v_n en fonction de n.

Exercice 64. On lance un dé équilibré à 6 faces jusqu'à obtenir deux six consécutifs. Le but de l'exercice est de calculer l'espérance du nombre de lancers nécessaires. On choisit de modéliser les lancers par une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes suivant la loi uniforme sur l'ensemble $\{1, \dots, 6\}$.

1. On pose $Z_0 = 0$, puis $Z_{n+1} = (Z_n + 1)\mathbb{1}_{\{X_{n+1}=6\}}$. Montrer que $(Z_0)_{n\geq 2}$ est une chaîne de Markov.
2. On pose $T = \inf\{n \geq 0; Z_n = 2\}$. Calculer $\mathbb{E}(T)$ par la méthode de l'analyse au premier pas. Conclure.

Exercice 65. Soit $(X_n)_{n\geq 0}$ une chaîne de Markov irréductible à valeurs dans un espace d'états fini ou dénombrable S. Soit $A \subset S$, avec A fini et $A \neq S$. On pose $\tau = \inf\{n \geq 0; X_n \notin A\}$. Montrer que $\tau < +\infty$ presque sûrement, et même qu'il existe $\beta > 0$ tel que $\mathbb{P}(\tau > k) = O(e^{-\beta k})$.

Exercice 66. *L'image d'une chaîne de Markov n'est pas (toujours) une chaîne de Markov.*
On considère la chaîne de Markov (X_n) sur $E = \{0, 1, 2\}$ de matrice de transition $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ et de loi initiale $\pi_0 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Soit $f : E \to \{0, 1\}$ telle que $f(0) = f(1) = 0$, $f(2) = 1$. Pour $n \geq 0$, on pose $Y_n = f(X_n)$. Montrer que $(Y_n)_{n\geq 1}$ n'est pas une chaîne de Markov.

Exercice 67. *Un collecteur de coupons biaisés.* [6]
Soit ν une probabilité sur $\{1,\dots,n\}$, où n est un entier avec $n \geq 2$. Soit $(X_k)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi ν. On pose $E_0 = \varnothing$, $Y_0 = 0$, puis pour $k \geq 1$, $E_k = \{X_1,\dots,X_k\}$ et $Y_k = |E_k|$. On peut imaginer que les X_k représentent la suite des lancers d'un dé à n faces, le dé n'étant pas nécessairement équilibré. Alors E_k est l'ensemble des faces qui sont sorties dans les k premiers coups, et Y_k le nombre de faces différentes apparues.

1. Montrer que (E_k) est une chaîne de Markov.
2. On suppose ici que ν est la loi uniforme sur $\{1,\dots,n\}$, c'est à dire que le dé est équilibré. Calculer $\mathbb{E}(Y_{k+1} - Y_k|E_1,\dots,E_k)$, puis $\mathbb{E}(Y_{k+1} - Y_k|Y_1,\dots,Y_k)$. Montrer alors que $(Y_k)_{k\geq 0}$ est une chaîne de Markov.
3. On suppose que $(Y_k)_{k\geq 0}$ est une chaîne de Markov. Exprimer $\mathbb{P}(Y_1 = 1, Y_2 = 1)$ et $\mathbb{P}(Y_1 = 1, Y_2 = 1, Y_3 = 1)$ en fonction des $\nu(k)$. En déduire que ν suit la loi uniforme sur un ensemble $J \subset \{1,\dots,n\}$.

Exercice 68. *L'image d'une chaîne de Markov peut être une chaîne de Markov.* [7]
Soit (X_n) une chaîne de Markov sur un ensemble dénombrable E de matrice de transition P. Soit ψ une application surjective de E dans un ensemble F telle que

$$\forall z \in F \quad \forall x,y \in E \quad \psi(x) = \psi(y) \Rightarrow \mathbb{P}^x(\psi(X_1) = z) = \mathbb{P}^y(\psi(X_1) = z).$$

1. Montrer que la suite (Y_n) définie par $Y_n = \psi(X_n)$ est une chaîne de Markov et déterminer sa matrice de transition.
2. On dit que μ est une probabilité stationnaire pour la dynamique de $(X_n)_{n\geq 0}$ réalisée sur $(\Omega,\mathcal{F},\mathbb{P})$ si on a $\mathbb{P}_{X_0} = \mathbb{P}_{X_1} = \mu$. Montrer que si μ est une probabilité stationnaire pour la dynamique de la chaîne (X_n) alors l'image de μ par ψ est une probabilité stationnaire pour la dynamique de la chaîne (Y_n).

6. Le titre est un clin d'œil à un modèle classique, le collecteur de coupons – non biaisé, celui là, qui sera étudié plus longuement dans l'exercice 100.

7. Ce résultat est parfois mentionné sous le nom de critère de Dynkin.

Exercice 69. *Urne avec remplacement*, extrait de Master Orléans 2005. Soit N un entier naturel non nul. Soient $(U_n)_{n\geq 1}$ et $(V_n)_{n\geq 1}$ deux suites de variables aléatoires indépendantes suivant la loi uniforme sur l'ensemble fini à N éléments $\{1,\dots,N\}$. On suppose également que les suites $(U_n)_{n\geq 1}$ et $(V_n)_{n\geq 1}$ sont indépendantes l'une de l'autre.

Pour tout $a \in \{0,\dots,N\}$, on peut définir par récurrence une suite $(X_n)_{n\geq 0}$ par

$$\begin{cases} X_0 = a \\ X_{n+1} = X_n + \mathbb{1}_{\{V_{n+1}\leq X_n\}} - \mathbb{1}_{\{U_{n+1}\leq X_n\}}. \end{cases}$$

1. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov à valeurs dans $\{0,\dots,N\}$ dont la matrice de passage $P = (p_{i,j})_{0\leq i,j\leq N}$ vérifie pour tout i dans $\{0,\dots,N\}$:

$$p_{i,i+1} = p_{i,i-1} = \frac{i(N-i)}{N^2} \text{ et } p_{i,i} = 1 - \frac{2i(N-i)}{N^2}.$$

   (Par convention $p_{0,-1} = p_{N,N+1} = 0$).
   On note désormais $\mathbb{P}^a$ la probabilité correspondante.

2. Pour $N = 5$, dessiner le graphe associé à cette chaîne de Markov.
3. Pour $i \in \{1,\dots,N-1\}$ et $n \in \mathbb{N}$, on pose

$$A_n^i = \{V_{n+1} \leq i\}\Delta\{U_{n+1} \leq i\}.$$

   Montrer que
$$\mathbb{P}(\overline{\lim}_{n\to+\infty}(A_n^i)) = 1.$$

4. Soit $i \in \{1,\dots,N-1\}$. Montrer que $\mathbb{P}(\lim_{n\to+\infty} X_n = i) = 0$.
5. Montrer que $(X_n)_{n\geq 0}$ est une martingale.
6. Montrer que X_n converge presque sûrement vers une variable aléatoire X_∞ qui prend ses valeurs dans $\{0, N\}$.
7. Déterminer la loi de X_∞. Indication : déterminer son espérance.

8. Une urne contient 20 boules noires et 10 boules blanches.

   On tire deux boules de l'urne, l'une après l'autre. On peint la deuxième des deux boules tirées avec la couleur de la première des deux boules tirées, puis on remet les deux boules dans l'urne. On répète alors sans cesse le même procédé. Quelle est la probabilité qu'au bout d'un certain moment, on ne tire plus que des boules noires ?

   Ce modèle est étudié en génétique sous le nom de *modèle de Moran*.

Exercice 70. *Chaîne de Bessel*, d'après Cachan 2007.
Soit $\Omega = \mathbb{Z}^{\mathbb{N}}$ et $\mathcal{F} = \mathcal{B}(\Omega)$. On note $(X_i)_{i\geq 0}$ les projections canoniques : pour $\omega \in \Omega$, on a $X_i(\omega) = \omega_i$. On peut alors définir une filtration $(\mathcal{F}_n)_{n\geq 0}$ en posant $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ pour tout $n \geq 0$.

Pour $x \in \mathbb{N}$, on note $\mathbb{P}^x$ la loi sous laquelle $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de passage $P = (p_{i,j})_{i,j\geq 0}$ vérifiant

$$\begin{cases} p_{k,k+1} = \frac{k+2}{2k+2} \text{ si } k \geq 1 \\ p_{k,k-1} = \frac{k}{2k+2} \text{ si } k \geq 1 \\ p_{0,0} = 1 \\ p_{i,j} = 0 \text{ dans tous les autres cas.} \end{cases}$$

et commençant au point x (c'est à dire que $\mathbb{P}^x(X_0 = x) = 1$). Pour $a \geq 0$, on pose $\tau_a = \inf\{k \geq 0; X_k = a\}$.

1. Soit f une fonction de $\mathbb{N}$ dans $\mathbb{R}$. Montrer que

   $$\mathbb{E}^x[f(X_{n+1})|\mathcal{F}_n] = \Psi_f(X_n),$$

   où

   $$\Psi_f(i) = \sum_{j=0}^{+\infty} p_{i,j} f(j).$$

2. On pose $Y_n = \frac{1}{1+X_n}$. Montrer que $(Y_n)_{n\geq 0}$ est une martingale bornée.
3. Soient a, b, x des entiers avec $0 \leq a < x < b$. On pose $\tau = \tau_a \wedge \tau_b$. Montrer que τ est un temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Dans la suite, on admettra que $\mathbb{P}^x(\tau < +\infty) = 1$ [8].

8. C'est une conséquence du résultat de l'exercice 65.

4. On pose $Z_n = Y_{n\wedge\tau}$. Montrer que $(Z_n)_{n\geq 0}$ est une martingale, puis, à l'aide du théorème de convergence dominée, montrer que

$$\mathbb{P}^x(\tau_a < \tau_b) = \mathbb{P}^x(Y_\tau = \frac{1}{1+a}) = \frac{\frac{1}{b+1} - \frac{1}{x+1}}{\frac{1}{b+1} - \frac{1}{a+1}}.$$

5. Montrer que $\lim_{n\to+\infty} \tau_n = +\infty$ $\mathbb{P}^x$ presque sûrement.
6. En déduire que $\mathbb{P}^x(\tau_a < +\infty) = \frac{a+1}{x+1}$.
7. On pose
$$I = \inf\{X_n; n \geq 0\}.$$
Montrer que sous $\mathbb{P}^x$, I suit la loi uniforme sur $\{0, \ldots, x\}$.

Exercice 71. *Une marche aléatoire sur le groupe affine Aff*$(\mathbb{F}_5)$. [9]
Le groupe des affinités sur $\mathbb{F}_5$ est formé des bijections de $\mathbb{F}_5$ dans lui-même qui s'écrivent $f(z) = az + b$, avec $(a, b) \in \mathbb{F}_5^\times \times \mathbb{F}_5$. C'est un sous-groupe de $\mathfrak{S}(\mathbb{F}_5)$. On définit une suite d'affinités aléatoires par $r_0 = \mathrm{Id}(\mathbb{F}_5)$, puis $r_{n+1} = f_{n+1} \circ r_n$, où $(f_n)_{n\geq 1}$ est une suite d'affinités aléatoires indépendantes de même loi, avec

$$\mathbb{P}(f_1 = g) = \mathbb{P}(f_1 = h) = \frac{1}{2}, \text{avec } g(z) = 2z \text{ et } h(z) = 3z + 1.$$

1. Montrer que $(r_n)_{n\geq 0}$ est une chaîne de Markov. Est-elle irréductible, apériodique ?
2. On cherche à calculer $\mathbb{P}(r_{2n} = \mathrm{Id})$ par une méthode matricielle.
   (a) On ordonne les états dans l'ordre

$$\begin{array}{llll} z \mapsto z, & z \mapsto z+1, & \ldots, & z \mapsto z+4, \\ z \mapsto 2z, & z \mapsto 2z+1, & \ldots, & z \mapsto 2z+4, \\ z \mapsto 3z, & z \mapsto 3z+1, & \ldots, & z \mapsto 3z+4, \\ z \mapsto 4z, & z \mapsto 4z+1, & \ldots, & \mapsto 4z+4. \end{array}$$

9. Pour d'autres résultats sur les marches aléatoires sur les groupes affines Aff$(\mathbb{F}_p)$, on pourra se rapporter à l'ouvrage de Diaconis [13], exemple 4 page 34. Cet ouvrage est librement téléchargeable à l'adresse https ://projecteuclid.org/euclid.lnms/1215467407

Vérifier que la matrice de la chaîne (r_n) est

$$M = \frac{1}{2}\begin{bmatrix} 0 & A & B & 0 \\ B & 0 & 0 & A \\ A & 0 & 0 & B \\ 0 & B & A & 0 \end{bmatrix},$$

avec

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \text{ et } B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

L'usage d'un outil numérique de calcul peut simplifier le travail.

(b) Montrer que (r_{2n}) est une chaîne de Markov. La représenter à l'aide de la matrice

$$N' = \begin{bmatrix} A' & B' \\ B' & A' \end{bmatrix},$$

où $A' = AB + BA$ et $B' = A^2 + B^2$.

(c) On admet que $N'^2 + N' = 2J_{10}$, où J_{10} est la matrice 10×10 dont toutes les entrées sont égales à 1. Montrer qu'il existe une suite $(\lambda_n)_{n\geq 1}$ telle que pour tout $n \geq 1$,

$$N'^n = \lambda_n J_{10} - (-1)^n N'.$$

(d) En déduire la valeur de $\mathbb{P}(r_{2n} = \text{Id})$

3. On cherche toujours à calculer $\mathbb{P}(r_{2n} = \text{Id})$, mais cette fois, on va utiliser des outils d'algèbre plus sophistiqués. On note G le sous-groupe de $\mathfrak{S}(\mathbb{F}_5)$ engendré par $t : z \mapsto z + 1$ et $s : z \mapsto -z$. G est le groupe diédral d'ordre 10. On admet les résultats d'algèbre suivants que l'on ne demande pas de redémontrer [10] :

$$\forall g \in G \quad \delta_{Id}(g) = \frac{1}{|G|} \sum_{\rho \in \text{Irr}(G)} \dim(\rho) \text{Tr } \rho(g), \tag{6.15}$$

10. Voir par exemple [32], corollaire 4.6 p 211 pour la formule (6.15) et la section 1.3 page 227 pour les résultats sur les groupes diédraux.

où $\mathrm{Irr}(G)$ désigne un système de représentants des représentations irréductibles de G. Lorsque $G = \mathbb{F}_5$, il y a 4 représentations irréductibles :

— 2 représentations de degré 1 : le caractère constant égal à 1 et le caractère envoyant t sur 1 et s sur -1.

— 2 représentations de degré 2 : si l'on pose $\omega = \exp(\frac{2i\pi}{5})$, on peut prendre les représentations qui envoient s sur $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et t respectivement sur $\begin{pmatrix} \omega & 0 \\ 0 & \omega^{-1} \end{pmatrix}$ et $\begin{pmatrix} \omega^2 & 0 \\ 0 & \omega^{-2} \end{pmatrix}$.

On rappelle qu'une représentation de degré k sur le groupe G est un morphisme de G dans $Gl_k(\mathbb{C})$.

(a) On pose $x_n = r_{2n}$. Montrer que $(x_n)_{n\geq 0}$ prend ses valeurs dans G et que

$$\mathbb{P}(x_1 = s) = \mathbb{P}(x_1 = t^{-1}s) = \mathbb{P}(x_1 = t) = \mathbb{P}(x_1 = t^2) = \frac{1}{4}.$$

(b) Montrer que pour tout $n \geq 0$,

$$\mathbb{P}(x_n = \mathrm{Id}) = \frac{1}{|G|} \sum_{\rho \in \mathrm{Irr}(G)} \dim(\rho) \mathrm{Tr}\ \left((\mathbb{E}(\rho(x_1)))^n \right).$$

Conclure.

Exercice 72. On pose $Y_0 = 0$, puis, pour $n \geq 1$, Y_n est une suite de variables indépendantes suivant la loi de Bernoulli de paramètre p. On pose ensuite $X_0 = 0$, et pour tout $n \geq 1$:

$$X_n = \max\{k : Y_n = Y_{n-1} = \dots Y_{n-k+1} = 1\} \wedge 0.$$

1. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov.
2. Dans la suite, on note $\mathbb{P}^i$ la loi d'une chaîne de Markov avec la même dynamique partant de n, $\mathbb{E}^i$ l'espérance correspondante. On note $T^n = \inf\{n \geq 0; X_i = n\}$. Montrer que pour $i < n$, on a

$$\mathbb{E}^i[T^n] = 1 + p\mathbb{E}^{i+1}[T^n] + (1-p)\mathbb{E}^0 T^n.$$

3. En déduire la valeur de $\mathbb{E}^0 T^n$.
4. Application : Une pièce de monnaie a pour probabilité p, de tomber sur face. On la lance indéfiniment. Calculer l'espérance du nombre de jets qu'il faudra jusqu'à ce qu'une chaîne de r résultats consécutifs de type face apparaisse.

Exercice 73. *Too big to fail.*
Le problème est le même que dans l'exercice 63, à ceci près que l'on suppose maintenant que la banque est infiniment riche [11]. On cherche la probabilité que le joueur soit ruiné.

Intuitivement, il suffit de faire tendre b vers $+\infty$ dans la formule (6.14), le tout étant de le justifier...

On suggère de poser pour tout k, $T_k = \inf\{n \geq 0; S_n = k\}$. Il s'agit donc de déterminer $\mathbb{P}(T_a < +\infty)$.

6.6.2 Exercices non corrigés

Exercice 74. *Chaîne à deux états.*
Soit $\{X_n : n \geq 0\}$ une chaîne de Markov à valeurs dans $\{0, 1\}$ et de probabilité de transition :

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \ 0 \leq \alpha, \beta \leq 1.$$

1. Montrer que pour $(\alpha, \beta) \neq (0, 0)$:

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

   Que se passe-t-il lorsque $\alpha = 0$ ou $\beta = 0$ ou $\alpha = \beta = 0$?

   On supposera pour la suite de l'exercice que $(\alpha, \beta) \neq (0, 0)$.
2. Vérifier que pour toute loi initiale μ, on a

   $$\mathbb{P}^\mu(X_n = 0) = \frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^n\left(\mu(0) - \frac{\beta}{\alpha+\beta}\right).$$

11. En cas de banqueroute, c'est le contribuable qui paye.

3. Si $(\alpha, \beta) \neq (1, 1)$, montrer que $(X_n)_{n\geq 0}$ converge en loi vers une loi ν que l'on déterminera. On supposera pour la suite de l'exercice que $(\alpha, \beta) \neq (1, 1)$.
4. (*Mesure stationnaire*) Prouver que, pour tout $n \in \mathbb{N}$,

$$\mathbb{P}^\nu(X_n \in A) = \nu(A).$$

Exercice 75. *Représentation canonique et simulation des chaînes de Markov.*

1. Soit $(Z_n)_{n\geq 1}$ une suite de vaiid à valeurs dans F, soit $g : E \times F \to E$ et soit X_0 une variable aléatoire à valeurs dans E indépendante de $(Z_n)_{n\geq 1}$. Montrer que la suite $(X_n)_{n\geq 0}$ définie par $X_{n+1} = g(X_n, Z_{n+1})$ est une chaîne de Markov homogène. Donner sa matrice de transition.
2. On suppose qu'on dispose d'un générateur de nombres aléatoires pour la loi uniforme sur $[0, 1]$, noté 'rand'. Soit μ une mesure de probabilité sur $\mathbb{N}$. Donner un algorithme pour générer des nombres aléatoires suivant la loi μ.
3. Soit $P = (p_{i,j})$ une matrice de transition sur $\mathbb{N}$.
On note $s_{i,k} = \sum_{j=0}^{k} p_{i,j}$. Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$, et X_0 une variable aléatoire à valeurs dans $\mathbb{N}$ indépendante de $(Z_n)_{n\geq 1}$. On construit la suite $(X_n)_{n\geq 0}$ par récurrence de la façon suivante :

$$\text{si } X_n(\omega) = i \text{ et } Z_{n+1}(\omega) \in]s_{i,j-1}, s_{i,j}] \text{ alors } X_{n+1} = j.$$

Montrer que la suite $(X_n)_{n\geq 0}$ ainsi définie est une chaîne de Markov homogène. Donner sa matrice de transition.
4. Application. Comment simuler une chaîne de Markov homogène de matrice de transition $P = (p_{i,j})$? Écrire un algorithme explicite si

$$P = \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}.$$

Exercice 76. *Temps d'atteinte d'un état absorbant.*

Soit (X_n) une chaîne de Markov sur un ensemble dénombrable E et $a \in E$ un état absorbant. On pose $T = \inf\{n \geq 0; X_n = a\}$. Montrer que $\mathbb{P}(X_n = a) = \mathbb{P}(T \leq n)$.

Exercice 77. *Temps d'entrée : une propriété d'invariance.*

Soit (X_n) une chaîne de Markov sur un ensemble dénombrable E de matrice de transition Q. Pour $f : E \to \mathbb{R}_+$, soit Qf la fonction définie par

$$Qf(x) = \sum_{y \in E} Q(x,y) f(y).$$

Pour $A \subset E$ on note $T_A = \inf\{n \geq 0; X_n \in A\}$ le temps d'entrée dans A. Montrer que la fonction f définie sur E par $f(x) = \mathbb{P}_x(T_A < +\infty)$ vérifie

$$f(x) = 1 \text{ pour } x \in A \text{ et } f(x) = (Qf)(x) \text{ pour } x \notin A.$$

Exercice 78. Soit $(X_n)_{n \geq 0}$ une chaîne de Markov à valeurs dans $\mathbb{N}$. On note A l'ensemble des points absorbant de la chaîne. Montrer que (X_n) ne peut converger que vers un élément de A.

Plus précisément : si il existe un événement B et une variable aléatoire Y telle que

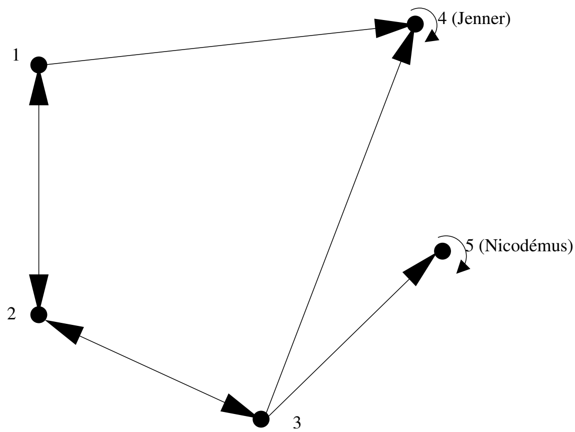
$$\forall \omega \in B \qquad \lim_{n \to +\infty} X_n(\omega) = Y(\omega),$$

alors $\mathbb{P}(B \cap \{Y \notin A\}) = 0$.

Exercice 79. *Madame Brisby dans le labyrinthe.* [12]

Madame Brisby s'est perdue dans le labyrinthe que forment les galeries où vivent les rats de Nimh. Quelle est la probabilité qu'elle rencontre le sage Nicodémus avant de croiser le belliqueux Jenner ?

12. *Madame Brisby et le secret de Nimh* est un long métrage d'animation américain réalisé par Don Bluth. Sorti en 1982, il est basé sur le roman de Robert C. O'Brien, *Madame Frisby et les Rats de Nimh*.



Exercice 80. Soit M la matrice d'une chaîne de Markov. Montrer que si $m_{i,i} > 0$, alors l'état i est apériodique. Qu'en déduire pour une chaîne irréductible ?

Exercice 81. Soit a et b des entiers supérieurs ou égaux à 2, $(D_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. à valeurs dans $\mathbb{Z}/a\mathbb{Z}\times\mathbb{Z}/b\mathbb{Z}$ vérifiant

$$P(D_1 = (0,1)) = P(D_1 = (1,0)) = \frac{1}{2}.$$

Soit $(D_n)_{n\geq 1}$ une suite de variables aléatoires et S_0 une variable aléatoire à valeurs dans $\mathbb{Z}/a\mathbb{Z}\times\mathbb{Z}/b\mathbb{Z}$ indépendante de $(D_n)_{n\geq 1}$ Pour $n \geq 1$, on pose

$$S_n = S_0 + \sum_{k=1}^{n} D_k.$$

Montrer que (S_n) est une chaîne de Markov . Est-elle irréductible, apériodique ?

Exercice 82. *Propriété de Markov fonctionnelle.*
Soit $(X_n)_{n\geq 1}$ une chaîne de Markov, F une application mesurable de $\mathbb{R}^{\mathbb{N}^*}$ dans $[0, +\infty[$.
Montrer que pour tout entier $p \geq 1$, et tout $A \in \mathcal{F}_p = \sigma(X_1, \dots, X_p)$, on a

$$\mathbb{E}[\mathbb{1}_A F((X_{n+p})_{n\geq 1})] = \mathbb{P}(A) g(X_p),$$

avec $g(x) = \mathbb{E}^x F((X_n)_{n\geq 1})$.

On rappelle que pour toute variable aléatoire Y positive et toute probabilité $\mathbb{P}$, on a $\mathbb{E}Y = \int_0^{+\infty} \mathbb{P}(Y > t)\, dt$.

Exercice 83. *Madame Brisby II.*
On reprend la chaîne de Markov des aventures de madame Brisby. On note l'espace d'états $E = \{1, 2, 3, 4, 5\}$ et l'on pose $A = \{4, 5\}$ Soit $f : E \to \mathbb{C}$ telle que $\forall x \in A \quad f(x) = 0$.

On pose $F = \sum_{k=1}^{+\infty} f(X_k)$.

Montrer que $\mathbb{E}|F| \leq (\mathbb{E}T - 1)\|f\|_\infty$.

Montrer l'identité

$$(I - N)\begin{pmatrix} \mathbb{E}^1 F \\ \mathbb{E}^2 F \\ \mathbb{E}^3 F \end{pmatrix} = \begin{pmatrix} \mathbb{E}^1 f(X_1) \\ \mathbb{E}^2 f(X_1) \\ \mathbb{E}^3 f(X_1) \end{pmatrix},$$

où N est la matrice 3×3 telle que la matrice de la chaîne de Markov admette une écriture par blocs sous la forme

$$\begin{pmatrix} N & * \\ 0 & I_2 \end{pmatrix}.$$

En déduire $\mathbb{E}^1 T, \mathbb{E}^2 T, \mathbb{E}^3 T$.

Exercice 84. *Évolution d'un génotype avec fixation : le modèle de Wright–Fisher.*
Nous travaillons ici sur une population de taille fixe formée de $2N$ gènes. Il y a deux types de gènes possibles : le type "a" et le type "A". Chacun des gènes au temps $n + 1$ est engendré par deux des $2N$ gènes présents au temps N. Son type est celui d'un de ses deux parents (choisi au hasard).

On considère la variable aléatoire X_n égale au nombre d'individus de type "A" dans la population à l'étape n.

On admettra qu'on peut modéliser l'évolution par la récurrence suivante :

$$X_{n+1} = \sum_{k=1}^{2N} \mathbb{1}_{\{Y_{n+1,k} \leq X_n\}},$$

où $(Y_{n,k})_{n\geq 1, k\in\{1,\dots,2N\}}$ est une suite de variables aléatoires indépendantes suivant la loi uniforme sur l'ensemble fini $\{1,\dots,2N\}$. X_0 est indépendante des $(Y_{n,k})$.

1. Montrer que X_n est une chaîne de Markov à valeurs dans l'ensemble $E = \{0,\dots,2N\}$.
2. Montrer que la loi de X_{n+1} sachant $X_n = k$ est une loi binomiale de paramètre $2N$ et $(k/2N)$. Identifier les éventuels points absorbants.
3. Montrer que $(X_n)_{n\geq 0}$ converge presque sûrement vers une variable aléatoire X_∞.
4. Déterminer la loi de X_∞ en fonction de la loi de X_0.

Exercice 85. Soit ν, μ deux lois sur $\mathbb{N}$. ν est appelée loi de reproduction et μ est la loi de la taille de la population initiale.

On appelle chaîne de Galton–Waltson de loi initiale μ et de loi de reproduction ν la chaîne de Markov de loi initiale μ et de matrice de transition

$$p_{i,j} = \begin{cases} \nu^{*i}(j) \text{ si } i \neq 0 \\ \delta_0(j) \text{ si } i = 0 \end{cases} .$$

Montrer que si $(X_n)_{n\geq 0}$ et $(Y_n)_{n\geq 0}$ sont deux chaînes de Markov indépendantes, $(X_n)_{n\geq 0}$ étant une chaîne de Galton–Waltson de loi initiale μ_1 et de loi de reproduction ν, et $(Y_n)_{n\geq 0}$ étant une chaîne de Galton-Waltson de loi initiale μ_2 et de loi de reproduction ν, alors $(X_n+Y_n)_{n\geq 0}$ est une chaîne de Galton–Waltson de loi initiale $\mu_1 * \mu_2$ et de loi de reproduction ν.