République Algérienne Démocratique et Populaire
الجمهورية الجزائرية الديمقراطية الشعبية
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
وزارة التعليم العالي و البحث العلمي

Université M'Hamed BOUGARA Boumerdès

Faculté des Sciences

Département des Mathématiques

Mémoire de fin d'études

Pour l'obtention du diplôme Master en Mathématique

Option : Mathématique Financière

Critères dérivées pour une chaîne de Markov image

Réalisé par :
AOUFI Imene
FEKHARI Asma

Encadré par :
Mme. LARABI : M.A.A (UMBB)

Soutenu le Juillet 2022, Devant le jury composé de :

Mme. DRICI	:	M.C.B	UMBB	- Présidente
Mme. IKHLEF	:	M.A.A	UMBB	- Examinateur

Année universitaire : 2021/2022

Remerciements

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux , qui nous a donné la force et la patience d'accomplir ce modeste travail.

*La première personne que nous tenons à remercier est notre encadrante "**Mme. LARABI Ghenima**" pour l'orientation , la confiance , la patience et les bonnes explications qui ont constitué un apport considérable sans lequel ce travail n'aurait pas pu être mené au bon port .*

*Nous tenons à exprimer notre profonde gratitude à "**Mme DRICI** " pour sa patience, sa confiance et sa gentillesse alors que nous préparons pour ce message. Elle était toujours disponible et s'est occupée de l'avancement de nos travaux jusqu'à la fin, ainsi que "**Mr. BENAMARA**" et "**Mme CHEMERIK**" qui nous ont aidés et ne nous ont pas épargnés.*

*Nos vifs remerciements vont également aux membres du jury "**Mme DRICI**" et "**Mme IKHLEF**" pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.*

Nous tenons à exprimer nos sincères remerciements à tous les professeurs qui nous ont enseigné et qui par leurs compétences nous ont soutenu dans la poursuite de nos études .

Enfin , nous tenons également à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail .

Dédicaces

Au terme de ce modeste travail , accompagné d'un profond amour , je dédie :

> “
>
> *À celle qui m'a toujours donné du courage et de la force , que j'ai toujours considéré comme mon modèle dans la vie ...ma chère mère.*
>
> *À celui qui était mon refuge et qui n'a rien épargné pour que j'arrive ici maintenant ... mon cher père.*
>
> *À mes frères, mes soeurs et mes baux frères que j'aime, et que j'espère voir dans les plus hauts rangs et postes à l'avenir.*
>
> *À toute ma famille du plus jeune au plus âgé et même à ceux qui nous ont quitté sans retour.*
>
> *À mes enseignants du primaire jusqu'à l'université.*
>
> *À ma bonne et sympathique collègue (binôme) " **Asma** " pour sa sérieusité et ses efforts.*
>
> *À tous mes collègues de la promotion 2021-2022 et à tous ce qui m'ont soutenu de prés ou de loin , même avec un mot d'encouragement ou une prière .*
>
> *Je dédie cet humble travail.*
>
> ”

- Imene

Dédicaces

“

À celle qui m’a toujours donné du courage et que nulle dédicace ne puisse exprimer ce que je lui dois . Qu’elle trouve dans ce travail la modeste récompense des sacrifices qu’elle a consentis pour moi... chère mère.

À l’homme qui m’a soutenu dans mes études. Aucun mot ne serait assez pour témoigner de l’étendue des sentiments que j’éprouve à son égard ...cher père.

À mes frères et ma soeur qui ont toujours été à mes côtés et avec qui j’ai passé des moments mémorables.

À toute ma famille pour leur soutien et même à ceux qui nous ont quitté sans retour.

À mes enseignants du primaire jusqu’à l’université .

À ma bonne et sympathique collègue (binôme) " ***Imene*** *" pour sa sérieusité et ses efforts.*

À tous mes collègues et à tous ce qui m’ont soutenu de prés ou de loin, même avec un mot d’encouragement ou une prière.

Je leur souhaite tant de réussite dans leur vie .

”

- Asma

Table des matières

Table des figures

Introduction générale

Les chaînes de Markov sont des outils mathématiques simples d'utilisation qui permettent de modéliser l'évolution au cours du temps (la dynamique) d'un système aléatoire. La propriété caractéristique d'une chaîne de Markov est que son évolution future ne dépend que du présent, pas du passé. On trouve des applications des chaînes de Markov dans beaucoup de domaines, de la biologie, à la gestion des ressources, en passant par les files d'attente des réseaux, la physique, les sciences de l'ingénieur, la recherche opérationnelle et les mathématiques financières.

Les chaînes de Markov apparaissent régulièrement en finance (modélisation de l'évolution de titres en bourse), en économie et dans bien d'autres domaines liés à la gestion. Nous pouvons facilement imaginer des structures comportant des pièces soumises à l'usure. Ces processus nous permettront de déterminer, dans plusieurs cas, la distribution des coûts de remplacement ainsi que les probabilités associées à des événements tels les moments de défaillance du système, etc.

L'objectif de ce mémoire est d'appliquer les critères dérivées d'une chaîne de Markov à une chaîne de Markov image.

Le premier chapitre aborde les concepts de base de la mesure et sa relation avec les probabilités (l'espace de probabilité , les applications mesurables) dont nous avons besoin pour fournir de connaissances sur notre sujet , et les notions des chaîne de Markov à temps discret avec différentes définitions, ses caractéristiques et son comportement asymptotique, ainsi les critères dérivés pour une chaîne de Markov.

Le deuxième chapitre forme le noyau de ce mémoire . Les outils de bases étant bien définis au cour du premier chapitre, on peut donc aborder "les critères dérivées d'une chaînes de Markov image".

Ce chapitre est scindé en deux parties :

• Dans la première partie on a vérifié sous certaines conditions que l'image d'une chaîne de Markov par une fonction est une chaîne de Markov et on a donné quelques résultat concernant la récurrence et la transience pour la chaîne de Markov image et aussi sur l'étude de comportement asymptotique d'une chaîne de Markov image.

• La deuxième partie, est l'objectif de notre travail . On a appliqué les critères dérivées sur une chaîne de Markov image et on a donné quelques résultats concernant cette partie.

Le troisième chapitre : Simulation des résultats précédents avec le programme Matlab afin de faciliter l'application numérique des théories étudiées.

Chapitre 1

Chaîne de Markov

Dans ce chapitre on a trois parties. La première aborde quelques concepts de base pour la notion de mesure et probabilité, la deuxième partie concerne les chaînes de Markov à temps discret, la troisième partie concerne les critères dérivés pour une chaîne de Markov.

1.1 Définitions

1.1.1 Ensembles

On appelle ensemble E toute collection d'objets, appelés éléments de l'ensemble E. Si le nombre de ces objets est fini, on l'appelle cardinal de E et on le note $Card(E)$, si E possède une infinité d'éléments, on dit quil est de cardinal infini et on note $Card(E) = \infty$. Si un objet x est un élément de E, on dit que x appartient à E et on note $x \in E$. Si x n'est pas un élément de E, on note $x \notin E$.

Il existe un ensemble, appelé l'ensemble vide et noté $\emptyset$, qui ne contient aucun élément. On a alors $Card(\emptyset) = 0$. [18]

Exemple : $\{0,1,3\}, \{rouge, bleue\}, \{0,1,2,3,...\} = \mathbb{N}$.

1.1.2 Partie d'un ensemble

On dit qu'un ensemble A est inclus dans un ensemble E, ou que A est une partie de l'ensemble E, ou que A est un sous ensemble de E si tout élément de A est un élément de E. On note $A \subset E$ et on a formellement :

$$A \subset E \Longleftrightarrow \forall x (x \in A \Rightarrow x \in E).$$

Quand A n'est pas une partie de E, on note $A \not\subset E$ et on a formellement :

$$A \not\subset E \Longleftrightarrow \exists x ((x \in A) \wedge (x \notin E)).$$

L'ensemble de toutes les parties d'un ensemble A est noté $\mathscr{P}(A)$.

Exemple :

Soient $E = \mathbb{N}$, A_1 le sous-ensemble formé par les entiers pairs, A_2 le sous-ensemble formé par les entiers impairs. Alors, les sous-ensembles A_1 et A_2 forment une partition de E.

1.1.3 Application

Soient E et F deux ensembles. On appelle application de E dans F un objet mathématique f qui à tout élément x de E associe un élément qu'on note $f(x)$ de F. Une telle application est notée

$$f : \left\{ \begin{array}{c} E \rightarrow F \\ x \mapsto f(x) \end{array} \right.$$

E et F s'appellent respectivement ensemble de départ et ensemble d'arrivée de f.
$f(x)$ est dite image de x par f, et x est appelée antécédent de $y = f(x)$. [15]

Exemple :

▷ La fonction f définie sur $\mathbb{R}$ par $f(x) = \sin x$, avec $F =]-1,1[$.
▷ L'application identique $idE = \mathbf{1}_E : E \rightarrow E$, définie par $idE(x) = \mathbf{1}_E(x) = x$.

1.1.4 Injective

On dit qu'une application $f : E \to F$ est injective ou que c'est une injection si l'une des propositions équivalentes suivantes est vraie :

$\triangleright \forall x, x' \in E,\ f(x) = f(x') \Rightarrow x = x'$.

$\triangleright \forall x, x' \in E,\ x \neq x' \Rightarrow f(x) \neq f(x')$.

$\triangleright$ Tout élément de F possède au plus un antécédent par f.

Les applications représentées sont injectives :[2]

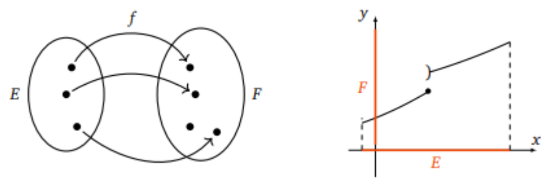

FIGURE 1.1 – Injective

1.1.5 Surjective

On dit qu'une application $f : E \to F$ est surjective ou que c'est une surjection si l'une des propositions équivalentes suivantes est vraie :

$\triangleright \forall y \in F,\ \exists x \in E,\ y = f(x)$.

$\triangleright \mathrm{Im} f = F$.

$\triangleright$ Tout élément de F possède au moins un antécédent par f.

Les applications représentées sont surjectives :

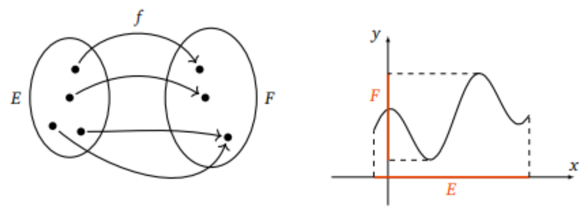

FIGURE 1.2 – Surjective

1.1.6 Bijective

On dit qu'une application $f : E \to F$ est bijective ou que c'est une bijection si l'une des propositions équivalentes suivantes est vraie.

$\triangleright \forall y \in F,\ \exists! x \in E,\ y = f(x)$.

$\triangleright f$ est injective et surjective.

$\triangleright$ Tout élément de F possède un unique antécédent par f.

Les applications représentées sont bijectives :

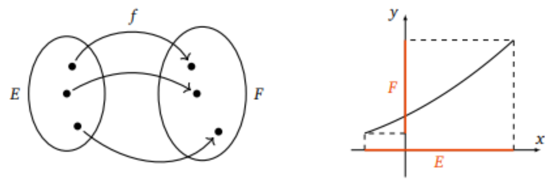

FIGURE 1.3 – Bijective

1.1.7 Image d'une application

Soit $f : E \to F$ une application. On appelle image de f l'ensemble noté Im f des éléments de F qui ont un antécédent par f dans E. Plus formellement

$$\operatorname{Im} f = \{y \in F \mid \exists x \in E,\ y = f(x)\} = \{f(x),\ x \in E\}.$$

1.1.8 Image directe

Soient $f : E \to F$ une application et A une partie de E. On appelle image directe de A par f notée $f(A)$, l'ensemble des éléments de F qui sont images d'élément de A (i.e. qui ont un antécédent dans A). Autrement dit,

$$f(A) = \{y \in F \mid \exists x \in A,\ y = f(x)\} = \{f(x),\ x \in A\}.$$

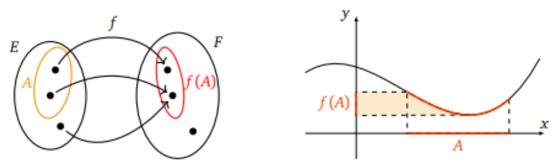

FIGURE 1.4 – Image directe d'une partie

1.1.9 Image réciproque

Soient $f : E \to F$ une application et B une partie de F. On appelle image réciproque de B par f, notée $f^{-1}(B)$, lensemble des éléments de E qui sont antécédents d'éléments de B (i.e. qui ont une image dans B). Autrement dit,

$$f^{-1}(B) = \{x \in E \mid f(x) \in B\}.$$

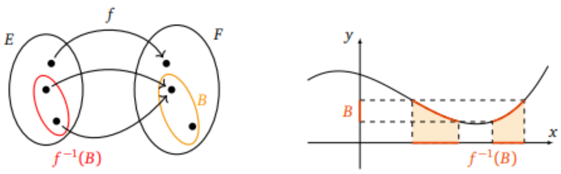

FIGURE 1.5 – Image réciproque d'une partie

1.1.10 Ensemble dénombrable

On dit un ensemble qu'il est dénombrable s'il est en bijection(injection si on considère les ensembles finis dénombrables) avec une partie de $\mathbb{N}$.

En particulier, un ensemble fini est considéré comme dénombrable. Certains auteurs définissent les ensembles dénombrables comme étant les ensemble en bijection avec $\mathbb{N}$.

1.1.11 Tribu

Soit E un ensemble quelconque. Une tribu (ou σ-algèbre) sur E est une famille $\mathscr{A}$ de parties de E telle que :
(i) $E \in \mathscr{A}$ et $\emptyset \in \mathscr{A}$.
(ii) A stable par passage au complémentaire, i.e

$$A \in \mathscr{A} \Rightarrow A^c \in \mathscr{A}.$$

(iii) A stable par réunion dénombrable, i.e

$$\forall n \in \mathbb{N}, A_n \in \mathscr{A} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}.$$

(iiii) A stable par intersection dénombrable, i.e

$$\forall n \in \mathbb{N} A_n \in \mathscr{A} \Rightarrow \bigcap_{n\in\mathbb{N}} A_n \in \mathscr{A}.$$

Les éléments de $\mathscr{A}$ sont appelés parties mesurables, ou parfois $\mathscr{A}$-mesurables s'il y a ambiguïté. On dit que $(E,\mathscr{A})$ est un espace mesurable. [12]

Exemple :
$\mathscr{A} = \mathscr{P}(E)$, On l'appelle tribu discrète.
$\mathscr{A} = \{\emptyset, E\}$, On l'appelle tribu trivial.

1.1.11.1 Tribu engendrée

Soit $\mathscr{C}$ un sous-ensemble de $\mathscr{P}(E)$. Il existe alors une plus petite tribu sur E qui contient $\mathscr{C}$. Cette tribu notée $\sigma(\mathscr{C})$ peut être définie par :

$$\sigma(\mathscr{C}) = \bigcap_{\mathscr{A}\ tribu,\ \mathscr{C}\subset\mathscr{A}} \mathscr{A}.$$

$\sigma(\mathscr{C})$ est appelée la tribu engendrée par $\mathscr{C}$.

Exemple : Tribu borélienne

La tribu des boréliens est la tribu sur $\mathbb{R}$ qui est engendrée par les intervalles de $\mathbb{R}$ (ouverts, fermes ou semi-ouverts),

$$\mathscr{B}r(\mathbb{R}) = \sigma(\{]a,b[, a,b \in \mathbb{R}\}).$$

1.1.11.2 Tribus image et image réciproque [1]

Soit $f : E \to F$

Proposition 1.1. *Si $\mathscr{B}$ est une tribu sur F,*

$$f^{-1}(\mathscr{B}) = \{f^{-1}(Y),\, Y \in \mathscr{B}\}.$$

est une tribu sur E, appelée tribu image réciproque (de $\mathscr{B}$ par f).

Preuve.
i) $f^{-1}(F) = E \in f^{-1}(\mathscr{B})$;
ii) pour tout $Y \in \mathscr{B}, (f^{-1}(Y))^c = f^{-1}(Y^c) \in f^{-1}(\mathscr{B})$;
iii) pour toute suite $(Y_n) \in \mathscr{B}, \bigcup_n f^{-1}(Y_n) = f^{-1}(\bigcup_n Y_n) \in f^{-1}(\mathscr{B})$ *car* $\bigcup_n Y_n \in \mathscr{B}$.

Proposition 1.2. *Si $\mathscr{A}$ est une tribu sur E,*

$$\mathscr{M} = \{Y \subseteq F :\ f^{-1}(Y) \in \mathscr{A}\}.$$

est une tribu sur F, appelée tribu image (de $\mathscr{A}$ par f).

1.1.12 Mesure positive

Définition 1.1. *[12]*

Une mesure positive sur $(E, \mathscr{A})$ est une application $\mu : \mathscr{A} \to [0, +\infty]$ qui vérifiée les propriétés suivantes :
(i) $\mu(\emptyset) = 0$;
(ii) Pour toute famille $(A_n)_{n\in\mathbb{N}}$ de parties mesurables disjointes deux à deux,

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

La propriété (ii) est appelée σ-additivité. Elle contient évidemment le cas particulier où les A_n sont vides à partir d'un certain rang, ce qui donne la propriété d'additivité finie.

On dit que $(E, \mathscr{A}, \mu)$ est un espace mesuré.

1- μ est dite finie si $\mu(E) < \infty$ (la quantité $\mu(E)$ est la masse totale de μ).

2- μ est une mesure de probabilité si $\mu(E) = 1$ et $(E, \mathscr{A})$ dit espace probabilisable.

3- μ est dite infinie si $\mu(E) = +\infty$ pour tout $A \in \mathscr{A}$.

1.1.13 Application mesurables

Définition 1.2. ***[11]***

Soient $(E,\mathscr{A})$ et $(F,\mathscr{B})$ deux espaces mesurables. Une application $f : E \to F$ est dite mesurable si et seulement si

$$\forall B \in \mathscr{B},\ f^{-1}(B) \in \mathscr{A}.$$

La composition de deux applications mesurables est encore mesurable. Cest immédiat en écrivant $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.

Proposition 1.3. *Pour que f soit mesurable, il suffit qu'il existe une sous-classe C de $\mathscr{B}$ telle que $\sigma(C) = \mathscr{B}$ et telle que la propriété $f^{-1}(B) \in \mathscr{A}$ soit vraie pour tout $B \in C$.*

Preuve. *Soit $\mathscr{G} = \{B \in \mathscr{B} : \ f^{-1}(B) \in \mathscr{A}\}$. Alors il est facile de vérifier que $\mathscr{G}$ est une tribu. Par hypothèse $C \subset \mathscr{G}$. Il en découle que $\mathscr{G}$ contient $\sigma(C) = \mathscr{B}$, d'où le résultat recherché.*

Définition 1.3. ***Mesure image [12]***

Soit $f : (E,\mathscr{A}) \to (F,\mathscr{B})$ une application mesurable, et soit μ une mesure positive sur $(E,\mathscr{A})$. La mesure-image de μ par f, notée $f(\mu)$ est la mesure positive sur $(F,\mathscr{B})$ définie par

$$f(\mu)(B) = \mu(f^{-1}(B)).$$

Il est facile de voir que la dernière formule définit bien une mesure sur $(F,\mathscr{B})$. Les mesures μ et $f(\mu)$ ont même masse totale, mais il peut arriver que μ soit σ-finie sans que $f(\mu)$ le soit.

1.1.14 Espace probabilisé

Définition 1.4. ***Probabilité [11]***

Soient $(E,\mathscr{A})$ un espace mesurable. On appelle probabilité sur $(E,\mathscr{A})$ toute mesure P sur $(E,\mathscr{A})$ de poids total 1, c'est à dire P est une application de $\mathscr{A}$ dans $[0,1]$ qui vérifie :
(i) $P(E) = 1$;
(ii) (Propriété de σ-additivité) Pour toute suite $(A_n)_{n\in\mathbb{N}^*}$ d'évènements, deux à deux disjoints, on a $P(\bigcup_{n\in\mathbb{N}^*} A_n) = \sum_{n=1}^{+\infty} P(A_n).$

On dit que le triplé $(E,\mathscr{A},P)$ est un espace de probabilité (probabilisé).

Remarque : Dans le cas générale on note Ω au lieu de E et P au lieu de μ pour les espaces probabilisés.

1.1.15 Probabilité conditionnelle

Soient A et B deux événements telle que $P(B) \neq 0$. On appelle Probabilité conditionnelle de A relativement à B ou de A sachant B, la probabilité que l'événement A se réalise sachant que B est réalisé. Cette probabilité vaut

$$P_B(A) = \frac{P(A\cap B)}{P(B)}.$$

Remarque : On trouve aussi la notation $P(A/B)$ pour $P_B(A)$.[4]

1.1.16 Variable aléatoire

Définition 1.5. *[24]*

Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé. Une variable aléatoire (v.a) à valeurs dans $(E, \mathscr{E})$ est une application $X : (\Omega, \mathscr{A}) \to (E, \mathscr{E})$ mesurable, c'est-à-dire telle que

$$\forall B \in \mathscr{E},\ X^{-1}(B) = \{X \in B\} = \{\omega \in \Omega :\ X(\omega) \in B\} \in \mathscr{A}.$$

On parle de variable aléatoire réelle(v.a.r) pour une v.a. à valeurs dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, et de vecteur aléatoire réel pour une v.a à valeurs dans $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$, avec $n \geq 2$.

De même une application $f : (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)) \to (\mathbb{R}^q, \mathscr{B}(\mathbb{R}^q))$ est borélienne si elle est mesurable pour les tribus des boréliens correspondantes :

$$\forall B \in \mathscr{B}(\mathbb{R}^q),\ f^{-1}(B) = \{x \in \mathbb{R}^n \mid f(x) \in B\} \in \mathscr{B}(\mathbb{R}^n).$$

▷ *Variable aléatoire discrète :C'est le cas où Ω est dénombrable (et $\mathscr{A}$ est l'ensemble des parties de Ω). La loi de X est alors*

$$P_X = \sum_{x \in E} P(X = x)\delta_x$$

autrement dit,

$$P_X(B) = \sum_{x \in B} P(X = x)$$

avec δ_x est la mesure de Dirac en x.
▷ *Variable aléatoire continue : Une variable aléatoire X à valeurs dans $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ est dite à densité si P_X est absolument continue par rapport à la mesure de Lebesgue λ telle que*

$$P_X(B) = \int_B p(x)dx$$

On note E l'espace d'états . On peut distinguer :
▷ *les espaces des états dénombrables, finis ou infinis ;*
▷ *les espaces des états continus.*

1.1.16.1 Espérance d'une variable aléatoire

Soit X une v.a à valeurs dans $(\Omega, \mathscr{A})$ et $f : E \to \mathbb{R}$ une application mesurable. On suppose que f est soit positive, soit intégrable par rapport à la loi de X (i.e $\int |f| dP_X < \infty$). On pose

$$\mathbb{E}[f(X)] = \int_E f(x) P_X(dx).$$

▷ Cas discrète :

$$\mathbb{E}[f(x)] = \sum_{x \in E} f(x) P(X = x).$$

▷ Cas réel à densité :

$$\mathbb{E}[f(X)] = \int_{-\infty}^{+\infty} f(x) p(x) dx.$$

Exemple :

1. On jette deux dés distincts et on s'intéresse à la somme des chiffres. On note X cette variable aléatoire, elle est définie par :

$$\begin{array}{rcl} X:\Omega & \rightarrow & \mathbb{R} \\ (\omega_1,\omega_2) & \mapsto & \omega_1+\omega_2. \end{array} \quad \text{avec} \quad \Omega=\{(1,1),(1,2),...,(6,5),(6,6)\}$$

L'ensemble des valeurs possibles de X est $\{2,3,...,12\}$.

2. On lance toujours deux dés, mais cette fois on s'intéresse au plus grand chiffre Y obtenu. On a alors

$$\begin{array}{rcl} Y:\Omega & \rightarrow & \mathbb{R} \\ (\omega_1,\omega_2) & \mapsto & \max(\omega_1,\omega_2). \end{array} \quad \text{avec} \quad \Omega=\{(1,1),(1,2),...,(6,5),(6,6)\}$$

La variable Y est à valeurs dans $\{1,2,...,6\}$.

1.1.17 Espérance conditionnelle

Définition 1.6. *[12]*

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé et X, Y deux variables aléatoires discrètes sur Ω. L'espérance conditionnelle de X sachant Y est la variable aléatoire réelle définie par

$$\mathbb{E}[X\mid Y]=\varphi(Y),$$

où la fonction $\varphi : E\rightarrow\mathbb{R}$ est donnée par

$$\varphi(y)=\begin{cases} \mathbb{E}[X\mid Y=y] & si\ P(Y=y)>0 \\ 0 & sinon \end{cases}$$

En particulier si X est également discrète,

$$\mathbb{E}[X\mid Y=y]=\sum_k x_kP(X=x_k\mid Y=y).$$

1.1.18 Processus stochastique (où Aléatoire)

On désigne par $(\Omega,\mathscr{A},P)$ un espace de probabilité.

Définition 1.7. *[24]*

Un processus stochastique (ou aléatoire) est une famille de variables ou de vecteurs aléatoires $(X_t)_{t\in T}$, indexée par l'ensemble $T\subset\mathbb{R}_+$ des temps dénombrable ou continu, définie sur un espace de probabilité $(\Omega,\mathscr{A},P)$ et à valeurs dans un espace d'états E.

$$X_t:\Omega\rightarrow E\,,\quad t\in T,$$

pour tout $\omega\in\Omega$ fixé, l'application :

$$t\mapsto X_t(\omega).$$

s'appelle une trajectoire du processus.
Des processus à temps discret ; on se limite alors à prendre $T\subset\mathbb{N}$; on voit que dans ce cas un processus stochastique revient à la donnée d'une suite de v.a.

1.1.19 Matrice stochastique

Soit $k \geq 1$ un entier et soit $P = (p_{i,j})_{(1 \leq i \leq k),\,(1 \leq j \leq k)}$ une matrice de taille $k \times k$ à coefficients réels. On dit que P est une matrice stochastique si et seulement si

$$\forall\, i, j \in \{1, ..., k\}, p_{i,j} \geq 0,$$

$$\forall\, i \in \{1, ..., k\}, \sum_{j=1}^{k} p_{i,j} = 1.$$

1.1.20 Noyau de transition

Définition 1.8.

Soient $(E, \mathscr{A})$ et $(F, \mathscr{B})$ deux espaces mesurables. On appelle Noyau de transition, de $(E, \mathscr{A})$ dans $(F, \mathscr{B})$ une application Q de $E \times \mathscr{B}$ dans [0, 1] telle que :
1. Pour tout $B \in \mathscr{B}$, l'application $x \to Q(x, B)$ est $\mathscr{A}$-mesurable,
2. Pour tout $x \in E$, l'application $B \to Q(x, B)$ est une probabilité sur $(F, \mathscr{B})$.[6]

1.2 Chaîne de Markov à temps discret

Soit $X_0, X_1, ..., X_n, ...$ une suite de variables aléatoires définies sur un même espace de probabilité $(\Omega, \mathscr{A}, P)$ et à valeurs dans l'espace d'état E.

1.2.1 Définition de chaîne de Markov à temps discret

Définition 1.9. *[23]*

La suite $X = (X_n)_{n \in \mathbb{N}}$ de variables aléatoires à valeurs dans E est une chaîne de Markov d'espace d'états E si pour tout $n \in \mathbb{N}$ et pour tout $i_1, ..., i_{n+1} \in E$ tel que $P(X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0) > 0$,

$$P(X_{n+1} = i_{n+1} \mid X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = P(X_{n+1} = i_{n+1} \mid X_n = i_n).$$

Définition 1.10. *[9]*

Une chaîne de Markov de noyau de transition $(P_n)_{n \geq 0}$ et de loi initiale μ_0 est une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ et à valeurs dans E telle que :

$$P(X_0 = i_0, ..., X_n = i_n) = \pi_0(i_0) \prod_{k=0}^{n-1} P(i_k, i_{k+1}).$$

pour tout $n \in \mathbb{N}^*$ et $i_0, ..., i_n \in E$.

Interprétation : La propriété de Markov signifie que l'état futur de la chaîne à l'instant $n+1$ dépend uniquement de son état à l'instant présent n et pas des instants passés par lesquels la chaîne est passée, On dit alors que le processus est "sans mémoire ".

1.2.2 La loi initiale de la chaîne de Markov

La loi initiale de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ à valeurs dans E est la loi μ_0 de X_0, tel que : $\mu_0 = P(X_0 = i), i \in E$. [24]

1.2.3 Chaîne de Markov homogène

Une chaîne de Markov (X_n) à valeurs dans un espace d'états E est dite homogène si

$$\forall\, n \geq 0, \forall\,(i,j) \in E^2, P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i).$$

Dans ce cas, on pose

$$p_{i,j} = P(X_1 = j \mid X_0 = i), i, j \in E.$$

P est la matrice de transition de la chaîne (X_n). [22]

1.2.4 Probabilités de transition

Soit $(X_n)_{n\geq 0}$ une chaîne de Markov homogène à valeurs dans un espace E. On appelle probabilités de transition la donnée, pour tout $n \geq 0$, pour tout $(i,j) \in E^2$

$$P[X_{n+1} = j \mid X_n = i] = p_{i,j} \quad (= P[X_1 = j \mid X_0 = i]).$$

1.2.5 Matrice de transition

Lorsque $E = \{1, ..., n\}$, la matrice $P = (p_{i,j})_{i,j\in E}$ est représentée par un tableau à n lignes et n colonnes appelée matrice de transition de la chaîne de Markov $(X_n)_n$:

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{pmatrix}$$

Remarque : La matrice de transition est une matrice stochastique.

1.2.6 Graphe de transition

A partir de la matrice de transition d'une chaîne de Markov, nous construisons un graphe. Il donne les mêmes informations que la matrice mais, comme toute représentation graphique, a l'avantage d'être plus parlant.

Définition 1.11. *[3]*
Étant donnée une chaîne de Markov $(X_n)_{n\geq 0}$ définie sur un espace d'états E, le graphe de la chaîne est le graphe construit à partir de la matrice de transition ainsi : les sommets sont les états et les arêtes (orientées) représentent les transitions possibles d'un état vers un autre. Au dessus de chaque arête on écrit la probabilité de transition correspondente.
De là, on obtient le graphe de transition suivant :

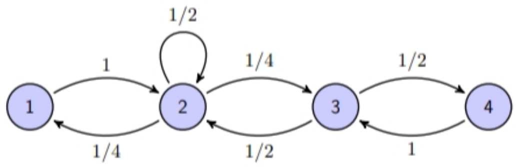

FIGURE 1.6 – Exemple d'un graphe de transition

1.2.7 Exemple d'une chaîne de Markov :

Disponibilité de deux machines : [20]

Une unité de production comporte deux machines qui fonctionnent indépendamment l'une de l'autre. Chaque machine fonctionne toute la journée avec la probabilité p ou bien tombe en panne durant la journée avec la probabilité $1-p$.

L'unité de production possède un technicien travailleur de nuit qui peut réparer une machine tombée en panne et la remettre en état de marche pour le lendemain. En revenche, le technicien ne peut réparer qu'une seule machine par nuit.

On souhaite comprendre le comportement du processus $\{X_n, n \geq 1\}$ où X_n représente le nombre de machines en panne au matin du n-ième jour.

Alors $\{X_n, n \geq 0\}$ est un processus stochastique à temps discret et à espace d'états $E = \{0, 1\}$ i.e :

$$\begin{cases} 0: & \text{aucune machine n'est en panne,} \\ 1: & \text{une machine est en panne.} \end{cases}$$

Signification des différentes probabilités de transition à calculer

$p_{00}(n)$: Aucune panne ou bien une panne est survenue;

$p_{10}(n)$: Une machine est tombée en panne et elle a été réparée la nuit, et la deuxième machine n'est pas tombée en panne;

$p_{01}(n)$: Les deux machines sont tombées en panne;

$p_{11}(n)$: L'autre machine est tombée en panne.

Aprés calcul, on aura :

$$\begin{aligned} p_{00}(n) &= P(X_{n+1} = 0 | X_n = 0) \\ &= p \cdot p + p \cdot (1-p) + (1-p) \cdot p \\ &= p^2 + 2p \cdot (1-p) \\ &= p \cdot (2-p). \end{aligned}$$

$$\begin{aligned} p_{10}(n) &= P(X_{n+1} = 0 | X_n = 1) \\ &= p. \end{aligned}$$

$$\begin{aligned} p_{01}(n) &= P(X_{n+1} = 1 | X_n = 0) \\ &= (1-p) \cdot (1-p) \\ &= (1-p)^2. \end{aligned}$$

$$\begin{aligned} p_{11}(n) &= P(X_{n+1} = 1 | X_n = 1) \\ &= (1-p). \end{aligned}$$

Le processus $\{X_n, n \geq 0\}$ est une chaîne de Markov,On vérifie par exemple :

$$\begin{aligned} P(X_{n+1}=0|X_n=1,X_{n-1}=1) &= P(X_{n+1}=0|X_n=1,X_{n-1}=0) \\ &= p \\ &= p_{10}(n). \end{aligned}$$

Donc le processus $\{X_n, n \geq 0\}$ est une chaîne de Markov. On a aussi $p_{i,j}(n) = p_{i,j} \quad \forall n \in \mathbb{N}$, Ce qui implique que $\{X_n, n \geq 0\}$ est une chaîne de Markov homogène dans le temps.
La matrice de transition associée :

$$P = \begin{pmatrix} p(2-p) & (1-p)^2 \\ p & (1-p) \end{pmatrix}$$

Le graphe de transition associé à la figure 1.7

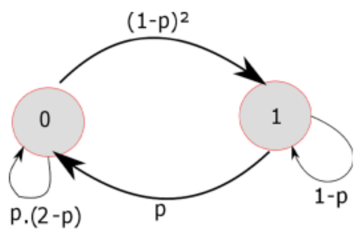

FIGURE 1.7 – Graphe de transition

1.2.8 Équation de Chapman-Kolmogorov

Pour tout $n \geq 0$, la matrice de transition en n coups est la puissance $n^{ème}$ de la matrice de transition de la chaîne, c'est-à-dire :

$$(p^n_{i,j})_{i,j} = P^n$$

Proposition 1.4. *[13]*

Pour tout $(i,j) \in E^2$ et tout couple (m,n) d'entiers positifs, on a l'identité :

$$P(X_{m+n} = j \mid X_0 = i) = \sum_{k \in E} P(X_m = k \mid X_0 = i) \cdot P(X_n = j \mid X_0 = k).$$

Ou encore

$$p^{(m+n)}_{i,j} = \sum_{k \in E} p^{(m)}_{i,k} \cdot p^{(n)}_{k,j}.$$

Preuve.

Pour $n = 0$, $P^{(0)} = I = P^0$(avec I est la matrice identité). Pour $n \geq 0$, on suppose que

$P^{(n)} = P^n$ et on veut montrer que $P^{(n+1)} = P^{n+1}$, c'est-à-dire que :

$$\begin{aligned}
p_{i,j}^{(n+1)} &= P(X_{n+1} = j \mid X_0 = i) \\
&= \sum_{k \in E} P(X_{n+1} = j, X_n = k \mid X_0 = i) \\
&= \sum_{k \in E} P(X_{n+1} = j \mid X_n = k, X_0 = i) \cdot P(X_n = k \mid X_0 = i) \\
&= \sum_{k \in E} P(X_{n+1} = j \mid X_n = k) \cdot P(X_n = k \mid X_0 = i) \\
&= \sum_{k \in E} p_{i,k}^{(n)} \cdot p_{k,j}.
\end{aligned}$$

ce qui signifie :

$$P^{(n+1)} = P^{(n)} \cdot P.$$

Donc :

$$P^{(n)} = P^n.$$

Exemple : *Soit $(X_n)_n \in \mathbb{N}$ une chaîne de Markov d'espace d'état $E = \{0, 1\}$ et de matrice de transition :*

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}$$

On calcule $p_{0,0}^{(2)}$
On a : Méthode 01 :On calcule :

$$\begin{aligned}
p_{0,0}^{(2)} &= \sum_{k=0}^{1} p_{0,k} \cdot p_{k,0} \\
&= p_{0,0} \cdot p_{0,0} + p_{0,1} \cdot p_{1,0} \\
&= 9/16 + 1/16 = 10/16.
\end{aligned}$$

Méthode 02 :On calcule

$$P^2 = P \cdot P = \begin{pmatrix} 10/16 & 6/16 \\ 6/16 & 10/16 \end{pmatrix}$$

La probabilité cherchée est :$p_{0,0}^{(2)} = 10/16$.

1.3 Classification des états

1.3.1 Communication

1.3.1.1 États accessibles [8]

Soient i et j deux états de E. On dit que l'état j est accessible depuis l'état i si

$$\exists n \in \mathbb{N}, p_{i,j}^{(n)} = P(X_n = j \mid X_0 = i) > 0.$$

1.3.1.2 États communiquants

On dit que les états i et j communiquent si chacun d'eux est accessible à partir de l'autre

$$i \longleftrightarrow j \Leftrightarrow \begin{cases} i \to j \\ j \to i \end{cases}$$

Pour que deux états ne communiquent pas il faut que l'un des deux ne soit pas accessible à partir de l'autre, c'est à dire que :

$$\forall n \geq 0 : p_{i,j}^{(n)} = 0 \quad \text{où} \quad \forall n \geq 0 : p_{j,i}^{(n)} = 0$$

Classes d'équivalence sur les états : Il est important de noter que la communication entre les états est une relation d'équivalence :
▷ **La réflexivité :** Tout état i de la chaîne communique avec lui même, i.e. $i \to i$.
▷ **La symétrie :** Si un état i communique avec un état j, alors la réciproque est vraie, i.e, $i \to j \Leftrightarrow j \to i$.
▷ **La transitivité :** Si un état i communique avec un état j qui lui même communique avec un état k, alors l'état i communique avec l'état k, i.e. si $i \to j$ et $j \to k$ alors $i \leftrightarrow k$.

1.3.2 Chaîne irréductible

Une chaîne de Markov pour laquelle il n'existe qu'une seule classe de communication (égale à l'ensemble des états) est dite irréductible.

Exemple : Considérons la chaîne de Markov dont l'espace des états est $E = \{1,2,3\}$, et de matrice de transition :

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{pmatrix}$$

et le graphe de transition associé à la chaîne X_n dans la figure 1.8 suivante :

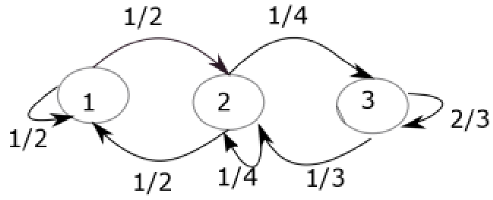

FIGURE 1.8 – Graphe de transition

Cette chaîne est irréductible, tous les états communiquent. Bien que $p_{1,3} = p_{3,1} = 0$ et qu'il n'y a donc pas de flèche entre les états 1 et 3, on a, en revanche, pour tout $n \geq 2$ et tout couple d'états (i, j) l'inégalité stricte : $p_{i,j}^{(n)} > 0$.

1.3.3 Récurrence

Quelques notations [19]

▷ Temps d'atteinte de i : $T_i = \inf\{n \geq 0 : X_n = i\}$,

▷ Nombre de passages par i : $N_i = \sum\limits_{n \geq 0} \mathbf{1}_{X_n = i} = Card\{n \geq 0 : X_n = i\}$,

▷ Nombre moyen de passages par i : $\mathbb{E}_j[N_i] = \sum\limits_{n \geq 0} P^{(n)}_{j,i}$,

▷ Temps de retour en i : $S_i = \inf\{n \geq 1 : X_n = i\}$,

▷ Temps moyenne de retour en i : $\mu_i = \mathbb{E}[T_i \mid X_0 = i] = \sum\limits_{n \geq 0} n \cdot p^{(n)}_{i,i}$.

▷ Nombre de retour en i : $L_i = \sum\limits_{n \geq 1} \mathbf{1}_{X_n = i} = \{n \geq 1 : X_n = i\}$.

1.3.3.1 État récurrent

Un état est dit récurrent si, partant de cet état, on est certain d'y revenir :

$$i \quad \text{récurrent} \Longleftrightarrow P_i(T_i < +\infty) = 1 \; et \; \sum_{n \in \mathbb{N}} p^{(n)}_{i,i} = +\infty.$$

1.3.3.2 État transient

Un état transient est un état pour lequel le retour n'est pas certain :

$$i \quad \text{transient} \Longleftrightarrow P_i(T_i < +\infty) < 1 \; et \; \sum_{n \in \mathbb{N}} p^{(n)}_{i,i} < +\infty.$$

1.3.3.3 Récurrent nulle et récurrent positive

Définition 1.12. *[16]*

Un point i de E est dit récurrent positif pour la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ si $\mathbb{E}_i[(T_i)] < +\infty$. Un point récurrent de E qui n'est pas récurrent positif est dit récurrent nul si $\mathbb{E}_i[(T_i)] = +\infty$.

Proposition 1.5. *Si E est fini, tous les états récurrents sont récurrents positifs.*

Remarque : Les chaînes de Markov irréductibles finies sont toujours récurrentes positives.

1.3.4 Chaîne absorbante

1.3.4.1 État absorbant

Définition 1.13. *[8]*

Un état est dit absorbant si, quand la chaîne atteint cet état, elle y reste :

$$i \quad absorbant \Longleftrightarrow \begin{cases} p_{i,i} = 1 \\ p_{i,j} = 0 \quad pour \quad j \neq i \end{cases}$$

Ces états sont très particuliers puisqu'ils constituent des états terminaux de l'évolution de la chaîne. Il est notamment intéressant d'étudier les probabilités d'adsorbtion, i.e. les probabilités que la chaîne finisse par atteindre un tel état.

Par définition, un état absorbant ne communique avec aucun autre : il constitue donc une classe (nécessairement récurrente) réduite à un seul élément.

Définition 1.14. *La chaîne de Markov est absorbante si*

$$\forall i, j \in E, \exists n \in \mathbb{N}^* : p_{i,j}^{(n)} \neq 0.$$

1.3.5 Chaîne régulière

On dit que la chaîne est régulière s'il existe un entier n tel que toutes les composantes de la matrice P^n sont strictement positives.

Exemple : La chaîne à deux états [14]

Notons $E = \{1, 2\}$ son espace d'états et sa matrice de transition est

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

avec p et q tous deux dans [0, 1].

Le graphe d'une telle chaîne est celui de la figure 1.9

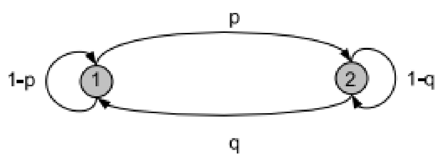

FIGURE 1.9 – Graphe de transition

La chaîne est :

▷ absorbante si $p = 0$ ou $q = 0$;

▷ irréductible non régulière si $p = q = 1$;

▷ régulière dans les autres cas.

1.3.6 Périodicité

1.3.6.1 Période d'un état

On définit la période d'un état i la quantité $d(i)$ telle que les retours à cet état se font seulement au bout de durées multiples de la période. L'état i est de période $d(i)$ si :

$$\{p_{i,i}^{(n)} > 0 \Rightarrow n = k\, d(i)\,, k \in \mathbb{N}\}$$

ou, plus précisément :

$$d(i)\,\text{période de}\, i \quad \Longleftrightarrow \quad d(i) = pgcd\{n \in \mathbb{N} : p_{i,i}^{(n)} > 0\}.$$

L'état i est dit périodique si $d(i) > 1$ et apériodique si $d(i) = 1$. [8]

Classe périodiques : La période est constante à l'intérieur d'une classe de communication. La période commune des éléments de la classe est appelée période de la classe.

1.3.7 Chaîne périodique

Une chaîne irréductible est réduite à une seule classe : si cette classe est périodique on parle de chaîne périodique. Si la chaîne n'a pas de période, elle est dite chaîne apériodique.

Exemple : Considérons la chaîne de Markov dont le graphe de transition est donné par la figure 1.10

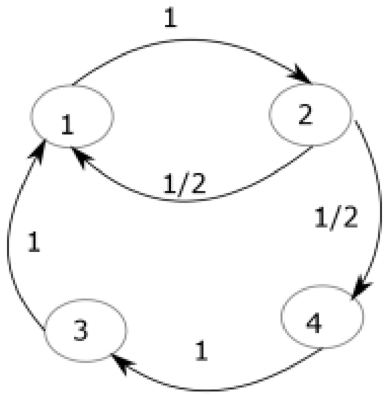

FIGURE 1.10 – Graphe de transition

Nous avons

$$\begin{aligned} d(1) &= pgcd\{n \geq 1, p_{11}^{(n)} > 0\} = \{2,4,6,8,10,\dots\}, \\ d(2) &= pgcd\{n \geq 1, p_{22}^{(n)} > 0\} = \{2,4,6,8,10,\dots\}, \\ d(3) &= pgcd\{n \geq 1, p_{33}^{(n)} > 0\} = \{4,6,8,10,12,\dots\}, \\ d(4) &= pgcd\{n \geq 1, p_{44}^{(n)} > 0\} = \{4,6,8,10,12,\dots\}. \end{aligned}$$

On constate que la période des états est 2.

1.4 Distribution stationnaire

Il est commande de noter les lois de probabilité π sur l'ensemble dénombrable E comme des vecteurs-lignes $\pi = (\pi_0, \pi_1, \ldots)$, où $\pi_{i \geq 0}$ et $\sum_i \pi_i = 1$.

Définition 1.15.

Soit π une loi de probabilité sur l'ensemble des états. Elle est dite une distribution stationnaire (ou invariante) si $\pi = \pi \cdot P$, ou de façon équivalente, si pour tout $j \in E$,on a :

$$\pi_j = \sum_{i \in E} \pi_i \, p_{i,j}.$$

On dit aussi "loi" où "probabilité invariante".

Remarque : Toute chaîne finie (nécessairement non transiente) admet au moins une distribution stationnaire, qui est unique si la chaîne est irréductible.[13]

1.4.1 Théorème d'existence dune distribution stationnaire

Théorème 1.1.

Si la matrice P a toutes ses éléments strictement positifs ou s'il existe une puissance P^n dont toutes les composantes sont strictement positives, alors il existe une distribution stationnaire π définie par le vecteur propre à gauche associé à la valeur propre $\lambda_1 = 1$. [8]

$$\pi = \pi \cdot P^n$$

1.4.2 Recherche des distributions stationnaires

Pour calculer les composantes du vecteur $\pi = (\pi_0, \pi_1, \cdots, \pi_n)$ d'une chaîne de Markov finie, on résout le système d'équations linéaire suivant : [20]

$$\begin{cases} \pi = \pi \cdot P \\ \sum_{k \in E} \pi_k = 1. \end{cases}$$

1.4.3 Probabilité réversible

On dit qu'une chaîne de Markov de matrice de transition P, ou plus simplement la matrice P, est réversible (ou symétrique) par rapport à la probabilité π si on a pour tous $i, j \in E$,

$$\pi_i p_{i,j} = \pi_j p_{j,i}.$$

Lemme 1.1. *Si une chaîne de Markov est réversible par rapport à la probabilité π, alors π est une probabilité invariante.[17]*

Exemple : Considérons la matrice de transition

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

d'une chaîne de Markov prenant ses valeurs dans l'ensemble $\{0, 1, 2, 3\}$ avec $\alpha \in [0,1]$

Recherchons les solutions stationnaires : La recherche de ces solutions stationnaires passe par la résolution du système suivant :

$$\begin{cases} \pi = \quad \pi.P \\ \sum_{k=0}^{3} \pi_k = 1 \end{cases} \Longleftrightarrow \begin{cases} \pi_0 = (1-\alpha)\pi_0 + \frac{1}{3}\pi_1 + \frac{1}{3}\pi_3 \cdots (1) \\ \pi_1 = \frac{1}{3}\pi_2 \cdots (2) \\ \pi_2 = \alpha\pi_0 + \frac{2}{3}\pi_1 + \frac{2}{3}\pi_3 \cdots (3) \\ \pi_3 = \frac{2}{3}\pi_2 \cdots (4) \\ \sum_{k=0}^{3} \pi_k = 1 \cdots (5) \end{cases}$$

Les équations 1) et 3) sont identiques d'où :

$$\begin{cases} \pi_1 = \frac{1}{3}\pi_2 \\ \pi_3 = \frac{2}{3}\pi_2 \\ \alpha\pi_0 = \pi_2 - \frac{2}{3}\pi_1 - \frac{2}{3}\pi_3 \quad = \frac{1}{3}\pi_2 \\ \sum_{k=0}^{3} \pi_k = 1 \end{cases}$$

On en déduit

$$\begin{cases} \pi_2 = 3\alpha\pi_0 \\ \pi_1 = \alpha\pi_0 \\ \pi_3 = 2\alpha\pi_0 \end{cases}$$

Cette dernière relation nous donne

$$\pi_0(1+\alpha+3\alpha+2\alpha) = 1$$

$$\pi_0(1+6\alpha) = 1$$

Par conséquent, la distribution stationnaire est donnée par :

$$\pi = \left(\frac{1}{1+6\alpha}, \frac{\alpha}{1+6\alpha}, \frac{3\alpha}{1+6\alpha}, \frac{2\alpha}{1+6\alpha}\right).$$

On notera que $\alpha \in [0,1]$ ce qui impliquera que $\alpha \neq \frac{-1}{6}$.

1.5 Distribution limite

Définition 1.16.

Une chaîne admet une distribution limite ou asymptotique π^*, si quelle que soit la loi initiale μ_0 :

$$\lim_{n\to+\infty} \pi(n) = \pi^*.$$

Plus explicitement, pour tout couple d'états (i,j), $\lim_{n\to+\infty} p_{i,j}^{(n)} =_{notée} \pi_j^*$ est indépendant de i.[25]

1.5.1 Calcul d'une distribution limite

Pour trouver la distribution limite π^* d'une chaîne de Markov, il faut diagonaliser la matrice P :

$$P = SDS^{-1}$$

- D : La matrice diagonale des valeurs propres.
- S : La matrice des vecteurs propres associés aux valeurs propres de P.
- S^{-1} : La matrice inverse de S.

1.5.2 Distribution stationnaire et distribution limite

Théorème 1.2. *Toute distribution limite π^* d'une chaîne de Markov, si elle existe, est l'unique distribution stationnaire π, nécessairement portée par les états récurrents.(la réciproque est fausse).[25]*

Exemple : Une unité de production comprend deux machines qui fonctionnent indépendamment l'une de l'autre. Chaque machine fonctionne avec probabilité $1/5$ au cours d'une journée, ce qui signifie que sa probabilité de tomber en panne pendant cette période est égale à $4/5$. Dans ce cas, elle sera réparée pendant la nuit et se retrouvera en état de marche le lendemain. Une seule machine péut-être réparer à la fois. Soit $X_n, n \geq 0$ le nombre de machines en panne au début de la n-ème journée. $p = 1/5, 1-p = 4/5$
Matrice de transition :

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

Aprés les calculs :

$$P = \begin{pmatrix} 9/25 & 16/25 \\ 5/25 & 20/25 \end{pmatrix}$$

Existance de la distribution limite
On a

$$P^2 = \begin{pmatrix} + & + \\ + & + \end{pmatrix}$$

Alors π^* existe et la chaîne converge vers π^*
Comme la chaîne est irréductible alors $\pi^* = \pi$

$$\begin{cases} \pi \cdot P = \pi \\ \sum_{i=0}^{1} \pi_i = 1 \end{cases}$$

$$\begin{cases} 9/25\pi_0 + 1/5\pi_1 = \pi_0 \cdots (1) \\ 16/25\pi_0 + 4/5\pi_1 = \pi_1 \cdots (2) \\ \pi_0 + \pi_1 = 1 \cdots (3) \end{cases}$$

De (1) on a :

$$16/25\pi_0 = 1/5\pi_1 \Rightarrow \pi_0 = 5/16\pi_1$$

On remplace π_0 dans (3) on obtient :

$$5/16\pi_1 + \pi_1 = 1 \Rightarrow \pi_1 = 16/21 \Rightarrow \pi_0 = 5/21$$

D'ou $\pi = (5/21, 16/21)$
La distribution limite $\pi^* = \pi = (5/21, 16/21)$
La matrice limite

$$\lim_{n \to +\infty} P^n = P^*$$

$$P^* = \begin{pmatrix} 5/21 & 16/21 \\ 5/21 & 16/21) \end{pmatrix}$$

1.6 Ergodicité

Définition 1.17.

Un état récurrent positif et apériodique est dit ergodique. Une chaîne irréductible, apériodique et récurrente positive est dite chaîne ergodique. [25]

1.6.1 Théorème ergodique

Théorème 1.3.

Soit X_n une chaîne de Markov homogène de loi initiale μ_0, irréductible et récurrente positive. On note π la probabilité invariante. Pour toute fonction $f \in L^1(\pi)$,

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow{n\to\infty} \pi \cdot f = \sum_{i \in E} \pi_i f(i), \quad P_\mu - p.s.$$

▷ *$f \in L^1(\pi)$ signifie que*

$$\sum_{i \in E} |f(i)|\pi_i < +\infty.$$

▷ *Si X_n est irréductible, pour tout $i \in E$, P_μ presque sûrement,*

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{X_k=i} \xrightarrow{n\to\infty} \frac{1}{\mathbb{E}_i[S_i]} = \begin{cases} \pi_i, & \text{si } X_n \text{ est récurrente positive,} \\ 0, & \text{sinon.} \end{cases}$$

▷ *En prenant, l'espérance, on obtient, pour tout $j \in E$ et $i \in E$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} p_{j,i}^{k} \xrightarrow{n\to\infty} \frac{1}{\mathbb{E}_i[S_i]} = \begin{cases} \pi_i, & \text{si } X_n \text{ récurrente positive,} \\ 0, & \text{si non.} \end{cases}$$

Proposition 1.6. *[19]*

Soit X_n une chaîne de Markov homogène, irréductible, récurrente positive et apériodique. Alors,

$$\forall i \in E, \forall j \in E, \lim_{n\to\infty} p_{j,i}^{(n)} = \pi_i.$$

▷ *Si X_n est irréductible, récurrente nulle alors,*

$$\forall i \in E, \forall j \in E, \lim_{n\to\infty} p_{j,i}^{(n)} = 0.$$

Théorème 1.4. *(de Chacon-Ornstein)*

Toute chaîne ergodique possède une distribution limite unique $\pi^* = (\pi_i^*)_i$ qui s'identifie à l'unique distribution stationnaire $\pi = (\pi_i)_i$ de la chaîne ;

$$\forall\, i, j \in E, \lim_{n \to +\infty} p_{j,i}^{(n)} = \pi_i^* = \pi_i = 1/\mu_i$$

où $\pi = \pi P$ et μ_i est le temps moyenne de retour en i. [25]

1.7 Martingale et chaîne de Markov

On considère la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ d'espace d'état E et de matrice de transition P.

Définition 1.18. *Une filtration de $(\Omega, \mathscr{A}, P)$ est une suite croissante $(\mathscr{F}_n)_{n\in\mathbb{N}}$ de sous-tribus de $\mathscr{A}$. On a donc*

$$\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots \subset \mathscr{A}$$

On dit aussi que $(\Omega, \mathscr{A}, (\mathscr{F}_n)_{n\in\mathbb{N}}, P)$ est un espace de probabilité filtré.
On interprète souvent le paramètre n comme un temps. La tribu $\mathscr{F}_n$ correspond alors à l'information acquise au temps n.

Exemple : *La filtration canonique associée à $(X_n)_n$ est définie par*

$$\mathscr{F}_n = \sigma(X_0, ..., X_n).$$

Définition 1.19. *[12]*
Soit $(X_n)_{n\in\mathbb{N}}$ un processus adapté à $\mathscr{F}_n$ si, pour tout $n \in \mathbb{N}$, X_n est $\mathscr{F}_n$-mesurable, tel que $\mathbb{E}[|X_n|] < \infty$.
On dit que le processus $(X_n)_{n\in\mathbb{N}}$ est :
▷ *une martingale si, pour tout $n \in \mathbb{N}$,*

$$\mathbb{E}[X_{n+1}|\mathscr{F}_n] = X_n;$$

▷ *une sur-martingale si, pour tout $n \in \mathbb{N}$,*

$$\mathbb{E}[X_{n+1}|\mathscr{F}_n] \leq X_n;$$

▷ *une sous-martingale si, pour tout $n \in \mathbb{N}$,*

$$\mathbb{E}[X_{n+1}|\mathscr{F}_n] \geq X_n;$$

1.7.1 Temps d'arrêt

Pour $i \in E$ on définit le temps aléatoire

$$T_i = \min\{n \geq 0 : X_n = i\},$$

premier moment où la chaîne atteint i. On étend cette définition à toute partie $A \subset E$:

$$T_A = \min\{n > 0 : X_n \in A\} = \min\{T_i : i \in A\}.$$

Définition 1.20. *[22]*
Soit X_n une chaîne de Markov, et T une variable aléatoire à valeurs dans $\mathbb{N}$.
T est un temps d'arrêt si pour tout n, l'évènement $(T = n)$ dépend uniquement du passé, c'est-a-dire si l'évènement $(T = n)$ est entièrement déterminé par les variables $X_0,...,X_n$.

Théorème 1.5. ***Probabilités et temps d'arrêts***
$(V_A(i))_{i\in E}$ et $(U_A(i))_{i\in E}$ sont solutions des systèmes (éventuellement infinis) déquations,si $A = \{j\}$:

$$V_A(i) = P(T_A < \infty) = \sum_{j\in E} P_{i,j}(V_A(j)), \text{pour tout } i \notin A$$

$$U_A(i) = \mathbb{E}_i(T_A) = 1 + \sum_{j\in E} P_{i,j}(U_A(j)), \text{pour tout } i \notin A$$

avec $V_A(i) = 1$ et $U_A(i) = 0$ pour $i \in A$.

1.7.2 Fonction harmonique

Définition 1.21. *[12]*

Une fonction $f : E \to \mathbb{R}^+$ est dite harmonique (resp. surharmonique) si on a pour tout $i \in E$, et P une matrice de transition d'une chaîne de Markov X_n,

$$f(i) = Pf(i)(resp.f(i) \geq Pf(i)).$$

Plus généralement, si $F \subset E$, on dit que f est harmonique sur F (resp. surharmonique sur F) si la propriété $f(i) = Pf(i)(resp.f(i) \geq Pf(i))$ est vraie pour $i \in F$.

Proposition 1.7. *[12]*

(i) La fonction f est harmonique (resp. surharmonique) ssi pour tout $i \in E$, le processus $(f(X_n))_{n\in\mathbb{N}}$ est une martingale (resp. une sur-martingale) sous P_i, relativement à la filtration $(\mathscr{F}_n)$.

(ii) Soit $F \subset E$ et $A = E\backslash F$. On note T_A le temps d'arrêt

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

Alors si f est harmonique (resp. surharmonique) sur F, le processus $(f(X_{n\wedge T_A}))_{n\in\mathbb{N}}$ est une martingale (resp. une sur-martingale) sous P_i, pour tout $i \in F$.

Preuve.

(i) Supposons d'abord f harmonique. Alors,

$$\mathbb{E}_i[f(X_{n+1}) \mid \mathscr{F}_n] = Pf(X_n) = f(X_n)$$

et en conséquence $\mathbb{E}_i[f(X_n) = \mathbb{E}_i[f(X_0) = f(i)$.
Inversement, supposons que $f(X_n)$ est une martingale sous P_i. Il vient immédiatement que

$$f(i) = \mathbb{E}_i[f(X_0) = \mathbb{E}_i[f(X_1) = Pf(i).$$

Le cas d'une fonction surharmonique est traité de la même façon.

(ii) Traitons le cas d'une fonction harmonique. On écrit pour $i \in F$

$$\begin{aligned}
\mathbb{E}_i[f(X_{(n+1)\wedge T_A}) \mid \mathscr{F}_n] &= \mathbb{E}_i[f(X_{n+1})\boldsymbol{1}_{(T_A>n)} \mid \mathscr{F}_n] + \mathbb{E}_i[f(X_{T_A})\boldsymbol{1}_{(T_A\leq n)} \mid \mathscr{F}_n] \\
&= \boldsymbol{1}_{(T_A>n)}\mathbb{E}_i[f(X_{n+1}) \mid \mathscr{F}_n] + f(X_{T_A})\boldsymbol{1}_{(T_A\leq n)} \\
&= \boldsymbol{1}_{(T_A>n)}Pf(X_n) + f(X_{T_A})\boldsymbol{1}_{(T_A\leq n)} \\
&= \boldsymbol{1}_{(T_A>n)}f(X_n) + f(X_{T_A})\boldsymbol{1}_{(T_A\leq n)} \\
&= f(X_{(n\wedge T_A)})
\end{aligned}$$

On a utilisé le fait que $f(X_{T_A})\boldsymbol{1}_{(T_A\leq n)} = f(X_{(T_A\wedge n)})\boldsymbol{1}_{(T_A\leq n)}$ est F_n-mesurable.

1.8 Exemple d'application d'une chaîne de Markov en finance

Le terme modèle de Markov s'est imposé dans les domaines de la finance, de l'économie et des sciences de gestion.
Le terme chaîne de Markov est souvent restreint aux modèles avec états et mesures dans un ensemble discret et à un instant discret. Cependant, il n'y a aucune raison pour que ces restrictions ne puissent pas être assouplies, pour inclure des observations d'une échelle continue.
La théorie Chaîne de Markov traite de l'estimation, qui comprend le filtrage du signal, la détermination des paramètres du modèle, l'estimation de l'état, le lissage du signal et la prédiction du signal, et le contrôle, qui fait référence à la sélection des actions qui affectent le système de génération de signal de manière à atteindre certains objectifs de contrôle .
Lors de la mise en uvre de Chaîne de Markov, des méthodes de probabilité standard sont utilisées. Il s'agit d'un ensemble de procédures conçues pour recadrer la tâche d'estimation et de contrôle d'origine dans un monde fictif afin que les résultats connus de variables aléatoires distribuées symétriquement et indépendamment puissent être appliqués.
Les résultats sont ensuite réinterprétés dans le monde réel en utilisant l'échelle de probabilité d'origine. Les cotes de crédit des entreprises et des gouvernements sont attribuées par des sociétés telles que CPA, BNA et CGCI ....
Les éléments étudiés lors du développement d'une entreprise ou d'un gouvernement sont la santé de son excédent, la qualité de ses actifs, la gestion du passif et de l'actif, le risque de réinvestissement et les flux monétaires futurs.
Bien sûr, un pointage de crédit élevé est plus attrayant qu'un pointage bas lorsqu'il s'agit d'investir dans l'une de ces sociétés. Une entreprise qui obtient une bonne cote de crédit est supposée présenter un faible risque.
Ainsi, les institutions financières lui prêteront de l'argent à un taux d'intérêt inférieur.

Exemple :

La firme BNA a choisi les cotes de crédit $AAA, AA, A, BBB, BB, B, CCC$. la cote AAA étant la meilleure. Il y a aussi la cote "defaut" qui est implicite et qui est attribuée lorsqu'une firme ne rencontre plus ses obligations.
Afin d'estimer les probabilités de passer d'une cote de crédit à une autre, il est possible de procéder comme suit :
la chaîne de Markov utilisée pour modéliser la variation de la cote de crédit d'une firme donnée au cours du temps sur un espace des états possibles.
Le processus stochastique $\{X_n : n \in \{0, 1, 2, ...\}\}$ représente donc la suite des cotes de crédit de la firme au cours du temps.
L'espace des états possibles est

$$E = \{x_1 = AAA, x_2 = AA, x_3 = A, x_4 = BBB, x_5 = BB, x_6 = B, x_7 = CCC, x_8 = Defaut\}$$

La matrice de transition de la chaîne de Markov définie antérieurement et qui sera associée à tous les problèmes de migration des cotes de crédit.
Le principal est de savoir ce que devra être une matrice carrée de dimension $[8 \times 8]$. Cette matrice devra contenir les probabilités de passer à l'une des huit cotes possibles étant donné l'une des huit possibilités de cotes actuelles.

La matrice de transition annuelle (car la période de temps choisie est d'une année) est estimée par

$$M=\begin{pmatrix} p_{AAA,AAA} & p_{AAA,AA} & p_{AAA,A} & \cdots & p_{AAA,CCC} & p_{AAA,Defaut}\\ p_{AA,AAA} & p_{AA,AA} & p_{AA,A} & \cdots & p_{AA,CCC} & p_{AAA,Defaut}\\ p_{A,AAA} & p_{A,AA} & p_{A,A} & \cdots & p_{A,CCC} & p_{A,Defaut}\\ p_{BBB,AAA} & p_{BBB,AA} & p_{BBB,A} & \cdots & p_{BBB,CCC} & p_{BBB,Defaut}\\ p_{BB,AAA} & p_{BB,AA} & p_{BB,A} & \cdots & p_{BB,CCC} & p_{BB,Defaut}\\ p_{B,AAA} & p_{B,AA} & p_{B,A} & \cdots & p_{B,CCC} & p_{B,Defaut}\\ p_{CCC,AAA} & p_{CCC,AA} & p_{CCC,A} & \cdots & p_{CCC,CCC} & p_{CCC,Defaut}\\ 0 & 0 & 0 & \cdots & 0 & p_{Defaut,Defaut} \end{pmatrix}$$

Le ratio $p_{AA,AA}=\frac{n_{AA,AAA}}{N_{AA}}$ est un estimé de la probabilité qu'une firme cotée AA au début de l'année soit cotée AAA à la fin de l'année (fin de la période).

• Classification des états de la chaîne de Markov est comme suit :

▷ Il existe deux classes possibles dans ce problème.

▷ La première étant une classe instable, elle posséde les états x_1 à x_7, soit les cotes $AAA, AA, \ldots, CCC$.

▷ La deuxième classe en est une stable car nous considérons ici qu'une compagnie qui est en defaut ne peut pas se redresser ce qui implique que l'état $Defaut$ est absorbant.

La cote $x_8 = Defaut$ est donc la seule cote appartenant à cette classe stable et elle est absorbante.

Comme la chaîne ne possède qu'un nombre fini d'états, les états de la classe instable sont transitoires tandis que $Defaut$ est un état récurrent.

▷ Comme tous les états d'une même classe ont la même période,alors la période est le plus grand commun diviseur de $\{1,2,3,\ldots\}$ qui est 1.

▷ En ce qui concerne l'état $Defaut$, le raisonnement est le même et sa période est aussi 1. Par conséquent, la chaîne est apériodique.

Pour plus d'informations sur l'utilisation d'une chaîne de Markov en finance consultez le document suivant "*Un modèle de migration de cotes de crédit utilisant les chaînes de Markov*".

1.9 Critéres dérivées pour une chaîne de Markov

1.9.1 Introduction

Les critères de dérivées encore appelés critères de fonctions de Lyapounov fournissent l'un des outils utilisés en pratique pour étudier le comportement en temps long des chaînes de Markov. Ils ont l'avantage de ne faire appel qu'à une analyse "à un pas" des transitions de la chaîne.

1.9.2 Un critère de non-récurrence positive

Théorème 1.6. *[10]*
Considérons un noyau de transition irréductible sur un ensemble fini ou dénombrable E, et supposons qu'il existe une fonction $V : E \to \mathbb{R}$ *vérifiant les propriétés suivantes :*
(1) pour tout $i \in E, V(i) \geq 0$;
(2) V *n'est pas constante sur* E;
(3) pour tout $i \in E, \mathbb{E}_i(V(X_1)) < +\infty$*;*
(4) pour tout $i \in E, \mathbb{E}_i(V(X_1) - V(X_0)) \geq 0$;
(5) il existe $c \geq 0$ *tel que, pour tout* $i \in E, \mathbb{E}_i \mid V(X_1) - V(X_0) \mid \leq c$.
Alors la chaîne est soit transiente, soit récurrente nulle.

Preuve. *Supposons la récurrence positive, et considérons* $i, j \in E$*, et* $T_1(i)$ *est un temps d'arrêt. On a*

$$V(X_{T_1(i)}) = V(X_0) + \sum_{k=0}^{T_1(i)-1} (V(X_{k+1}) - V(X_k)),$$

soit

$$V(X_{T_1(i)}) = V(X_0) + \sum_{k=0}^{+\infty} (V(X_{k+1}) - V(X_k)) \mathbf{1}_{(T_1(i)>k)}.$$

On voit alors facilement que

$$\sum_{k=0}^{+\infty} \mathbb{E}_j(\mid V(X_{k+1}) - V(X_k) \mid \mathbf{1}_{(T_1(i)>k)}) \leq c \sum_{k=0}^{+\infty} P_j(T_1(i) > k) < +\infty,$$

en utilisant (5) et la récurrence positive, on a donc

$$\mathbb{E}_j \sum_{k=0}^{+\infty} (V(X_{k+1}) - V(X_k)) \mathbf{1}_{(T_1(i)>k)} = \sum_{k=0}^{+\infty} \mathbb{E}_j((V(X_{k+1}) - V(X_k)) \mathbf{1}_{(T_1(i)>k)}),$$

et le fait que

$$\mathbb{E}_j(V(X_{k+1}) - V(X_k) \mid \mathscr{F}_k) \geq 0$$

d'après (4) entraîne que

$$\mathbb{E}_j((V(X_{k+1}) - V(X_k)) \mathbf{1}_{(T_1(i)>k)}) \geq 0,$$

car $T_1(i) > k$ *est un événement de* $\mathscr{F}_k$*. (*$\mathscr{F}_k$*est la tribu engendrée par les variables aléatoires* $\sigma(X_1, \cdots, X_n)$*) Or, avec probabilité 1 sous* P_j*, on a* $V(X_0) = V(j)$*, et, en utilisant la récurrence supposée,* $V(X_{T_1(i)}) = V(i)$*.*
On déduit donc de cette positivité le fait que $V(i) \geq V(j)$ *pour tous* i *et* j*, donc* V *doit être constante.Contradiction.*

1.9.3 Un critère de transience

Théorème 1.7. *[10]*
Considérons un noyau de transition irréductible sur un ensemble fini ou dénombrable E, et supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes
(1) pour tout $i \in E, V(i) \geq 0$;
(2) V n'est pas constante sur E ;
(3) pour tout $i \in E, \mathbb{E}_i(V(X_1)) < +\infty$;
(4) pour tout $i \in E$, $\mathbb{E}_i(V(X_1) - V(X_0)) \leq 0$;
Alors la chaîne est transiente. La réciproque est vraie.

Preuve. *On remarque que les hypothèses entraînent automatiquement que*

$$\forall n, j, \mathbb{E}_j(V(X_n)) < +\infty.$$

Ce résultat a en fait déjà été vu : V est une fonction surharmonique positive non-constante sous P_i, ce qui est impossible si la chaîne est récurrente.

Pour la réciproque, on suppose qu'il y a transience, et l'on fixe un point a et on prend la fonction définie par $V(i) = P_i(T_1(a) < +\infty)$ si $i \neq a$ et $V(a) = 1$.

1.9.4 Un critère de récurrence

Théorème 1.8. *[10]*
Considérons un noyau de transition irréductible sur un ensemble fini ou dénombrable E, et supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) pour tout $i \in E, \mathbb{E}_i(V(X_1)) < +\infty$;
(3) il existe un sous-ensemble fini C de E, tel que, pour tout $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq 0;$$

(4) pour tout K, l'ensemble $\{i \in E, V(i) \leq K\}$ est fini.
Alors la chaîne est récurrente.

Preuve. *Considérons $j \in E \setminus C$. On remarque que l'hypothèse entraîne facilement que*

$$\forall n : \mathbb{E}_j(V(X_n)) < +\infty.$$

Posons $T_1(C) = \inf\{n \geq 0; X_n \in C\}$ (avec la convention habituelle $inf\,\phi = +\infty$).
On pose

$$Y_n = V(X_n)\mathbf{1}_{(T_1(C)>n)}.$$

En utilisant la positivité de V , on obtient que, pour tout $n \geq 0$,

$$\mathbb{E}_j(V(X_{n+1})\mathbf{1}_{(T_1(C)>n+1)} \mid \mathscr{F}_n) \leq \mathbb{E}_j(V(X_{n+1})\mathbf{1}_{(T_1(C)>n)} \mid \mathscr{F}_n).$$

Comme $T_1(C)$ est un temps d'arrêt, l'événement $\{T_1(C) > n\}$ est mesurable par rapport à $\mathscr{F}_n$, et l'on a donc

$$\mathbb{E}_j(V(X_{n+1})\mathbf{1}_{(T_1(C)>n)} \mid \mathscr{F}_n) = \mathbb{E}_j(V(X_{n+1})) \mid \mathscr{F}_n)\mathbf{1}_{(T_1(C)>n)}.$$

Comme, compte-tenu de l'hypothèse (3), on a, sur l'évènement $X_n \notin C$, l'inégalité

$$\mathbb{E}_j(V(X_{n+1})) \mid \mathscr{F}_n) \leq V(X_n),$$

on en déduit finalement que

$$\mathbb{E}_j(Y_{n+1} \mid \mathscr{F}_n) \leq Y_n,$$

et donc le fait que (Y_n) est une surmartingale positive par rapport à $(\mathscr{F}_n)$, pour la probabilité P_j.
Par le théorème de convergence des surmartingales, Y_n tend donc avec probabilité 1 (sous P_j) vers une limite (a priori aléatoire) finie lorsque $n \to \infty$, et donc, sur l'événement $\{T_1(C) = +\infty\}, V(X_n)$ tend avec probabilité 1 vers une limite finie.
Par ailleurs, si nous supposons le noyau transient, l'hypothèse (4) entraîne que, pour tout entier k l'ensemble, fini d'après nos hypothèses, $\{i \in E; V(i) \leq k\}$ n'est visité qu'un nombre fini de fois par la chaîne, et par conséquent on doit avoir $\lim\limits_{n\to+\infty} V(X_n) = +\infty$ avec probabilité 1. En revenant à (Y_n), on constate donc que l'événement $\{T_1(C) = +\infty\}$ doit donc avoir une probabilité nulle sous P_j, j pouvant être choisi arbitrairement hors de l'ensemble C. Par irréductibilité de la chaîne, et en utilisant le fait que C est un ensemble fini, on en déduit facilement que la chaîne est récurrente, d'où une contradiction. On conclut donc finalement que p est transient.

1.9.5 Un critère de récurrence positive

Théorème 1.9. ***(Critère de Foster)[10]***
Considérons un noyau de transition irréductible sur un ensemble fini ou dénombrable E, et supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :
(1) pour tout $i \in E, V(i) \geq 0$;
(2) pour tout $i \in E, \mathbb{E}_i(V(X_1)) < +\infty$;
(3) il existe $\varepsilon > 0$ et un sous-ensemble fini C de E, tel que, pour tout $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq -\varepsilon.$$

Alors la chaîne est récurrente positive. La réciproque est vraie.

Preuve. *On remarque que l'hypothèse entraîne automatiquement que $\mathbb{E}_j(V(X_n)) < +\infty$ pour tous j et n. On reprend les notations utilisées dans la preuve du théorème précédent. D'après celui-ci, on sait déjà qu'il y a récurrence. On vérifie facilement que les hypothèses entraînent le fait que, pour $i \notin C$,*

$$\mathbb{E}_i(Y_{n+1}) \leq \mathbb{E}_i(Y_n) - \varepsilon \cdot P_{i_{(T_1(C) > n)}}.$$

On en déduit facilement en itérant et en utilisant la positivité de Y , que

$$\mathbb{E}_i(Y_0) \geq \varepsilon \sum_{k=0}^{n} P_i(T_1(C) > k),$$

d'où le fait que

$$\mathbb{E}_i(T_1(C)) \leq \frac{V(i)}{\varepsilon}.$$

Notons que, pour $i \in C$, on peut écrire que

$$\mathbb{E}_i(T_1(C)) = 1 + \sum_{j \notin C} p(i,j) \mathbb{E}_j(T_1(C)),$$

d'où

$$\mathbb{E}_i(T_1(C)) \leq 1 + \sum_{j \notin C} \frac{p(i,j)V(j)}{\varepsilon} \leq 1 + \sum_{j \in E} \frac{p(i,j)V(j)}{\varepsilon},$$

la dernière expression étant finie grâce à l'hypothèse (2). On obtient ainsi une borne explicite sur $\mathbb{E}_i(T_1(C))$ en termes de V et ε.
On conclut ensuite à la récurrence positive de la façon suivante. Partant de $j \in C$, définissons une suite de temps aléatoires par $\tau_0 = 0$ et, pour tout

$$k \geq 0, \tau_{k+1} = \inf\{n \geq \tau_k; X_n \in C\}.$$

On vérifie facilement que les τ_i sont des temps d'arrêt, et que la suite $(X_{\tau_k})_{k\geq 0}$ est une chaîne de Markov sur l'ensemble C, irréductible car la chaîne de départ l'est, et donc positivement récurrente du fait que C est fini. Pour $i \in C$, appelons

$$H(j) = \inf\{k \geq 1; X_{\tau_k} = j\}.$$

Nous savons donc que $\mathbb{E}_j(H_j) < +\infty$. Pour tout $k \geq 0$, posons $E_k = \tau_{k+1} - \tau_k$. Notons que l'on a l'identité

$$T_1(j) = \tau_{H(j)} = \sum_{k=0}^{H(j)} E_k = \sum_{k=0}^{+\infty} E_k \mathbf{1}(H(j) > k).$$

A présent, notons que l'événement $H(j) > k$ est mesurable par rapport à la tribu $\mathscr{F}_{\tau_k}$. Par conséquent, la propriété forte de Markov appliquée à l'instant τ_k entraîne que

$$\mathbb{E}_i(E_k \mathbf{1}(H(j) > k)) \leq M P_x(H(j) > k),$$

où

$$M = \sup_{i \in C} \mathbb{E}_i(T_1(C)),$$

en utilisant le fait que $X_{\tau_k} \in C$. D'après ce qui précède, $\mathbb{E}_i(T_1(C)) < +\infty$ pour tout $i \in E$, et C est un ensemble fini, donc $M < +\infty$. On en déduit que

$$\mathbb{E}_i(T_{\mathbf{1}}(j)) \leq M \cdot \sum_{k=0}^{+\infty} P_i(H(j) > k) = M \cdot \mathbb{E}_j(H(j)) < +\infty.$$

Pour la réciproque, on prend $a \in E$, et l'on pose $V(i) = \mathbb{E}_x(T_1(a))$ pour $i \neq a, V(a) = 0,$ et $C = \{a\}$.

1.9.5.1 Un critère d'ergodicité géométrique

Théorème 1.10. *[10]*
Considérons un noyau de transition irréductible sur un ensemble fini ou dénombrable E, et supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes
(1) pour tout $i \in E, V(i) \geq 1$;
(2) pour tout $i \in E, \mathbb{E}_i(V(X_1)) < +\infty$;
(3) il existe $0 < \varepsilon$ et un sous-ensemble fini C de E, tel que, pour tout $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq -\varepsilon V(i).$$

Alors il existe $i \in E, a, b > 0$ tels que, tout $n \geq 0$,

$$P_i(T_1(i) \geq n) \leq a \cdot \exp(-bn).$$

Preuve. *Notons que la positivité de V impose que $\varepsilon \leq 1$. On reprend les notations précédentes, en posant cette fois*

$$Y_n = \rho_0^n V(X_n) \mathbf{1}(T_1(C) > n),$$

où $\rho_0 > 1$ tel que $\rho_0(1-\varepsilon) \leq 1$. On vérifie que, pour $i \notin C$,

$$\mathbb{E}_i(Y_{n+1} \mid \mathscr{F}_n) \leq \rho_0(1-\varepsilon) Y_n,$$

et, par conséquent, pour tout $n \geq 0$,

$$\mathbb{E}_i(Y_n) \leq \mathbb{E}_i(Y_0) \leq V(i).$$

En utilisant l'hypothèse (1), on en déduit que, pour tout $n \geq 0$,

$$\rho_0^n P_i(T_1(C) > n) \leq V(i),$$

d'où, pour tout $\rho < \rho_0$, une inégalité de la forme

$$\mathbb{E}_i(\rho T_1(C)) \leq cV(i),$$

où $c \geq 0$ dépend de ρ_0 et ρ, mais pas de i.
A présent, pour tout $i \in C$, on a

$$\mathbb{E}_i(\rho T_1(C)) = \rho \sum_{j \notin C} p(i,j) \mathbb{E}_j(\rho T_1(C)) + \sum_{j \in C} p(i,j) \rho$$

d'où

$$\mathbb{E}_i(\rho T_1(C)) \leq \rho + \sum_{j \notin C} p(i,j) \mathbb{E}_j(\rho T_1(C)) \leq \rho + c\, \mathbb{E}_i(V(X_1)) < +\infty.$$

On reprend la décomposition utilisée dans la preuve du théorème précédent :

$$T_1(j) = \tau_{H(j)} = \sum_{k=0}^{H(j)} E_k.$$

En utilisant le fait que $(X_{\tau_k})_{k \geq 0}$ est une chaîne de Markov sur un ensemble fini, on déduit que, pour $j \in C, H(j)$ possède une queue sous-géométrique sous P_j. En suite, pour tout k, la propriété forte de Markov entraîne que

$$\mathbb{E}_j(\rho E_k \mid \mathscr{F}_{\tau_k}) = \mathbb{E}_{X_{\tau_k}}(\rho^{T_1(C)}).$$

Or, en vertu de ce qui précède, C étant un ensemble fini,

$$\sup_{i \in C} \mathbb{E}_i(\rho T_1(C)) < +\infty.$$

Au vu de ce résultat, et des estimations précédentes, on constate $(E_k)_{k \geq 0}$.
A présent, pour tout $\alpha > 0$, nous pouvons écrire que, par la borne de la réunion,

$$P_j(T_1(j) \geq n) \leq P_j(H(j) \geq \alpha n) + P_j\left(\sum_{k=0}^{\lfloor \alpha n \rfloor} E_k \geq n\right).$$

Comme $H(j)$ possède une queue sous-géométrique, $P_j(H(j) \geq \alpha n)$ est exponentiellement petit en n pour tout $\alpha\big(\sum_{k=0}^{|\alpha n|} E_k \geq n\big)$ est exponentiellement petit en n pourvu que α soit suffisament petit. On en déduit que, pour tout $j \in E$, il existe $a, b > 0$ tels que $P_j(T_1(j) \geq n) \leq a \cdot exp(-bn)$, ce qui prouve le résultat.

Pour prouver la réciproque, on considère un point a tel que $T_1(a)$ possède une queue sous-géométrique sous P_a. On utilise le théorème d'ergodicité géométrique, qui entraîne facilement l'existence de $r > 1$ tel que $\mathbb{E}_i(rT_1(a)) < +\infty$ pour tout $i \in E$. On pose alors $V(i) = \mathbb{E}_i(rT_1(a))$, pour $i \notin a, V(a) = 1$, et $C = \{a\}$.

Chapitre 2

Critères dérivées pour une chaîne de Markov image

Dans ce chapitre on a deux parties. La première aborde les conditions pour lesquelles l'image d'une chaîne de Markov image soit une chaîne de Markov par la fonction f et l'étude des propriétés de $f(X_n)$ relativement aux propriétés de X_n , la deuxième partie concerne l' application des critères dérivés d'une chaîne de Markov sur une chaîne de Markov image avec des conditions sur la fonction f (injective(bijective) et surjective) et sur la matrice de transition .

2.1 Image d'une chaîne de Markov

Le but de cette partie est de trouver une réponse à la question suivante : L'image d'une chaîne de Markov par une fonction est-elle une chaîne de Markov ?

Dans le cas général, la réponse est non, pour le prouver nous allons traiter un contre-exemple.

Exemple : Soit $(X_n)_{n\in\mathbb{N}}$ est la marche aléatoire simple issue de 0 et $f(X_n) = \mathbf{1}_{X_n>0}$ Alors : $(Y_n)_{n\in\mathbb{N}} = (f(X_n))_{n\in\mathbb{N}}$ n'est pas une chaîne de Markov, car :

$$P[Y_2 = 1/Y_1 = 1, Y_0 = 1] = 1/2.$$

$$P[Y_2 = 1/Y_1 = 1] = 1.$$

2.2 Chaîne de Markov image

Dans le cas général l'image d'une variable aléatoire par une fonction f est toujours une variable aléatoire. Par contre, l'image d'une chaîne de Markov par une fonction f n'est pas une chaîne de Markov.

Dans la suite, nous allons tenter de savoir s'il existe des cas où cette propriété est vérifiée ? Nous allons traiter les différentes applications particulières : injective , surjective et non bijective.

2.3 Cas d'une application injective (bijective)

Proposition 2.1.

Soit $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov, et $f : E \to F$ une fonction injective(bijective) telle que $Y_n = f(X_n)$. Alors $f(X_n)$ est une chaîne de Markov.

Autrement dit, la injection (bijection) de f est une condition suffisante pour que l'image d'une chaîne de Markov soit une chaîne de Markov.

Preuve. *Soient $y_0, y_1, ..., y_n, y_{n+1} \in F$. On a $Y_n = f(X_n)$,*
Donc

$$\begin{aligned}
P(Y_{n+1} = y_{n+1}/Y_n = y_n, \cdots, Y_0 = y_0) &= P(X_{n+1} = f^{-1}(y_{n+1})/X_n = f^{-1}(y_n), \cdots, X_0 = f^{-1}(y_0)) \cdots (1) \\
&= P(X_{n+1} = f^{-1}(y_{n+1})/X_n = f^{-1}(y_n)) \cdots (2) \\
&= P(Y_{n+1} = y_{n+1}/Y_n = y_n).
\end{aligned}$$

(1) :car f est injective(bijective).

(2) :car (X_n) est une chaîne de Markov.

Proposition 2.2. *(condition suffisante)*

Si $f : E \to F$ est injective (bijective) et $Y_n = f(X_n)$, alors (Y_n) est une chaîne de Markov de matrice de transition

$$Q = q_{x,y} = p_{f^{-1}(x),f^{-1}(y)}$$

pour tout x et y dans F tel que $f^{-1}(x)$ et $f^{-1}(y)$ existent.

De plus si $(X_n)_n$ possède une distribution stationnaire $(\pi_i)_{i\in E}$ alors $(Y_n)_n$ possède une distribution stationnaire $\pi' = (\pi'_y)_{y\in F}$ tel que

$$\begin{cases} \pi'_y = \pi_{f^{-1}(y)} & si\ f^{-1}(y)\ existe; \\ \pi'_y = 0 & sinon. \end{cases}$$

Preuve. $\pi' = (\pi)_{y\in F}$ *est une distribution stationnaire ssi* $\sum\limits_{y\in F} \pi_y = 1$ *et* $\sum\limits_{x\in F} \pi'_x q_{x,y} = \pi_y$.

$\triangleright \sum\limits_{y\in F} \pi_y = \sum\limits_{y\in F} \pi_f^{-1}(y) = \sum\limits_{i\in E} \pi_i = 1.$

$\triangleright \sum\limits_{x\in F} \pi'_x q_{x,y} = \sum\limits_{x\in F} \pi_{f^{-1}(x)} p_{f^{-1}(x),f^{-1}(y)} = \sum\limits_{i\in E} \pi_i p_{i,j} = \pi_j = \pi_{f^{-1}(y)} = \pi'_y$

Exemple : Soit X_n une chaîne de Markov d'espace d'états E, de distribution stationnaire $\pi = (1/3, 1/2, 1/6)$ et de matrice de transition

$$P = \begin{pmatrix} 1/4 & 3/4 & 0 \\ 1/2 & 1/3 & 1/6 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

et de graphe de transition

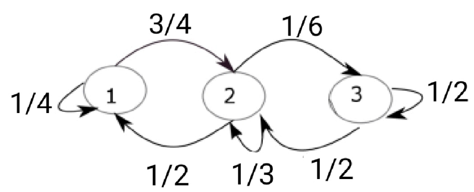

FIGURE 2.1 – Graphe de transition

Soit $f : E = \{1,2,3\} \longrightarrow F = \{1,2,3,4\}$ une application injective(bijective) tel que :

$$f(1) = 3, f(2) = 1, f(3) = 2.$$

$$\begin{cases} q_{x,y} = p_{f^{-1}(x)f^{-1}(y)} \\ q_{11} = p_{f^{-1}(1)f^{-1}(1)} = p_{22} = 1/3 \\ q_{12} = p_{f^{-1}(1)f^{-1}(2)} = p_{23} = 1/6 \\ q_{13} = p_{f^{-1}(1)f^{-1}(3)} = p_{21} = 1/2 \\ q_{22} = p_{f^{-1}(2)f^{-1}(2)} = p_{33} = 1/2 \\ q_{21} = p_{f^{-1}(2)f^{-1}(1)} = p_{32} = 1/2 \\ q_{23} = p_{f^{-1}(2)f^{-1}(3)} = p_{31} = 0 \\ q_{33} = p_{f^{-1}(3)f^{-1}(3)} = p_{11} = 1/4 \\ q_{31} = p_{f^{-1}(3)f^{-1}(1)} = p_{12} = 3/4 \\ q_{32} = p_{f^{-1}(3)f^{-1}(2)} = p_{13} = 0 \end{cases}$$

$$Q = \begin{pmatrix} 1/3 & 1/6 & 1/2 \\ 1/2 & 1/2 & 0 \\ 1/4 & 3/4 & 0 \end{pmatrix}$$

La matrice Q est une matrice stochastique.
La distribution stationnaire image est :
$\pi' = (\pi'_y)$ avec $\pi'_y = \pi_{f^{-1}(y)}$
$\pi'_1 = \pi_{f^{-1}(1)} = \pi_2 = 1/2$
$\pi'_2 = \pi_{f^{-1}(2)} = \pi_3 = 1/6$
$\pi'_3 = \pi_{f^{-1}(3)} = \pi_1 = 1/3$

$$\pi' = (1/2, 1/6, 1/3).$$

2.3.1 Caractérisation d'une chaîne de Markov image

Proposition 2.3.

Si (X_n) est irréductible et $f : E \to F$ une application injective(bijective) tel que

$$Y_n = f(X_n) \quad avec \quad P(Y_n = y_n, .., Y_0 = y_0) \neq 0,$$

alors la chaîne de Markov $(f(X_n))_{n\geq 0}$ est irréductible.

Preuve. *Soit $f : E \to F$ une application injective(bijective)*

Montrons que si $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov irréductible, alors $Y_n = (f(X_n))_{n\geq 0}$ est irréductible.
Soient

$$\forall x, y \in F, \exists i, j \in E \quad tel\ que \quad x = f(i),\ y = f(j)$$

(X_n) Une chaîne de Markov irréductible

$$\forall i, j \in E, \exists n, m \in \mathbb{N}^* \quad tel\ que \quad p_{i,j}^{(n)} > 0 \quad et \quad p_{i,j}^{(m)} > 0.$$

(Y_n) Une chaîne de Markov Image irréductible

$$\forall x, y \in F, \exists n, m \in \mathbb{N}^* \quad tel\ que \quad q_{x,y}^{(n)} > 0 \quad et \quad q_{x,y}^{(m)} > 0.$$

Soient $x, y \in F$, $q_{x,y} = p_{f^{-1}(x), f^{-1}(y)}$
Donc :

$$q_{x,y}^{(n)} = p_{f^{-1}(x),f^{-1}(y)}^{(n)} = p_{i,j}^{(n)} > 0 \quad \textit{(car } (X_n) \textit{ irréductible)};$$

$$q_{x,y}^{(m)} = p_{f^{-1}(x),f^{-1}(y)}^{(m)} = p_{i,j}^{(m)} > 0 \quad \textit{(car } (X_n) \textit{ irréductible)}.$$

D'où la chaîne $(Y_n)_n$ *est irréductible.*

Proposition 2.4.

Soit $(X_n)_{n\in\mathbb{N}}$ *une chaîne de Markov homogène d'espace d'état* E *et* $f : E \to F$ *une application injective (bijective).*

Alors, pour $i \in E$ *et* $x \in F$*, tel que* $x = f(i)$ *:*
1) i *état récurent* $\Rightarrow$ x *récurent.*
2) i *état transitoire* $\Rightarrow$ x *transitoire.*
3) i *état récurent positive* $\Rightarrow$ x *récurent positive.*

Preuve. *Soit* $\forall\, i \in E, \exists\, x \in F$ *tel que* $f(i) = x$
1) $\sum_{n\geq 1} q_{x,x}(n) = \sum_{n\geq 1} p_{f^{-1}(x),f^{-1}(x)}(n) = \sum_{n\geq 1} p_{i,i}(n) = \infty$ *(car i état récurrent) alors x récurrent.*
2) $\sum_{n\geq 1} q_{x,x}(n) = \sum_{n\geq 1} p_{f^{-1}(x),f^{-1}(x)}(n) = \sum_{n\geq 1} p_{i,i}(n) < \infty$ *(car i état transitoire) alors x transitoire.*
3) $\mathbb{E}_x(T_x) = \sum_{n\geq 0} n \cdot P(T_x = n) = \sum_{n\geq 0} n \cdot p^n_{f^{-1}(x),f^{-1}(x)} = \sum_{n\geq 0} n \cdot p^n_{i,i} < \infty$ *(car i état récurrent positive)*

Alors x récurrent positive.

Proposition 2.5.

Soit $(X_n)_{n\in\mathbb{N}}$ *une chaîne de Markov homogène,* $f : E \to F$ *une fonction injective(bijective), Soit* $i \in E$, $x = f(i) \in F$ *alors :*
1) i *apériodique* $\Rightarrow$ x *apériodique.*
2) i *périodique* $\Rightarrow$ x *périodique.*

Preuve. *soit* $i \in E, x = f(i) \in F$

$$d(x) = PGCD\{n \geq 1; q^n_{x,x} > 0\} = PGCD\{n \geq 1; p^n_{f^{-1}(x),f^{-1}(x)} > 0\} = d(i).$$

Alors
$\triangleright\, d(i) = 1((X_n), est\ apériodique) \Longleftrightarrow d(x) = 1 \Rightarrow x\ est\ apériodique.$
$\triangleright\, d(i) > 1((X_n), est\ périodique) \Longleftrightarrow d(x) > 1 \Rightarrow x\ est\ périodique.$

2.3.2 Le comportement asymptotique d'une chaîne de Markov image

Soient $(X_n)_n$ une chaîne de Markov d'espace d'états E fini, de distribution limite π^* et f une application injective(bijective) tel que $Y_n = f(X_n)$ une chaîne de Markov image d'espace d'états F et de distribution limite $\pi' = (\pi'_y)_{y\in F}$ tel que $\pi'_y = \pi_{f^{-1}(y)}$.

2.3.2.1 Chaîne de Markov irréductible apériodique

Proposition 2.6.

Si $(X_n)_n$ est irréductible apériodique alors la chaîne $(Y_n)_n$ est irréductible apériodique.

$$\forall x,y \in F, \lim_{n\to+\infty} q_{x,y}^{(n)} = \lim_{n\to+\infty} p^n_{f^{-1}(x),f^{-1}(y)} = \pi'.$$

Si la chaîne $(Y_n)_n$ est ergodique on a :

$$\lim_{n\to=\infty} Q^{(n)} = Q^*.$$

Où Q^* est une matrice dans tout les lignes égales à π'.

2.3.2.2 Chaîne de Markov irréductible périodique

Proposition 2.7.

Si $(Y_n)_n$ est irréductible périodique (de période $d > 1$).
1) On fait une partition de F en classe $C_1, C_2, .., C_d$ disjointes tel que il existe une transition se passe d'une classe C_k à C_{k+1}, $k = 1, ..., d$.
2) En numérotant de manière consécutive les états de chaque classe, on obtient une matrice diagonale R par blocs qui correspondent à sous chaînes irréductibles et apériodiques tel que :

$$R = (r_{x,y})_{x,y\in F} = Q^{(d)} = (q_{x,y}^{(d)})_{x,y\in F}$$

$$R = \begin{pmatrix} q_1 & 0 & \cdots & \cdots & 0 \\ 0 & q_2 & \cdots & \cdots & 0 \\ \vdots & \vdots & q_3 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & q_d \end{pmatrix}$$

Proposition 2.8.

Soit $(X_n)_n$ une chaîne de Markov irréductible périodique alors $(Y_n)_n$ est irréductible périodique (de période $d > 1$).

$$\lim_{n\to+\infty} r_{x,y}^{(n)} = \lim_{n\to\infty} p^{(nd)}_{f^{-1}(x),f^{-1}(y)} = \begin{cases} d\pi' & \text{si } f^{-1}(x), f^{-1}(y) \text{ appartient à même classe ;} \\ 0 & \text{si non} \end{cases}$$

$$r_{x,y}^{(n)} = (q^d_{x,y})^{(n)} = q_{x,y}^{(nd)} = p^{(nd)}_{f^{-1}(x),f^{-1}(y)}$$

On a :

$$\lim_{n\to+\infty} \sum_{k=1}^{n} \pi_k = \frac{1}{d} \sum_{i=1}^{d} \pi^{C_i}$$

2.4 Cas d'une application surjective

Dans cette partie, nous allons aborder le cas d'une application " non injective ", et donc non bijective.
Si $f : E \to F$ est surjective tel que : $Y_n = f(X_n)$. On ne peut rien dire sur le caractère markovien de la suite (Y_n) car dans le cas général, la surjection n'assure pas cette propriété.

Proposition 2.9.

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov. Soit $f : E \to F$ une application mesurable surjective telle que $Y_n = f(X_n)$.
On suppose que $\forall y \in F, \forall x, x' \in E$:

$$f(x) = f(x') \to P_x(f(X_1) = y) = P_{x'}(f(X_1) = y)$$

Alors $(Y_n)_n$ est une chaîne de Markov.

Proposition 2.10. *(condition suffisante)*

Soient $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène et $f : E \to F$ une application mesurable surjective telle que : $Y_n = f(X_n)$.
Si $\forall\, i, j \in E$:

$$f(i) = f(j) \Rightarrow \sum_{k\in E/f(k)=x} p(i,k) = \sum_{k'\in E/f(k')=x} p(j,k').$$

Alors $(Y_n)_n$ est une chaîne de Markov de probabilité de transition

$$q_{x,y} = \sum_{j\in E/f(j)=y} p_{i,j} \quad \textit{telle que} \quad f(i) = x,\ x \in F.$$

Proposition 2.11.

Si π est une distribution stationnaire pour la suite (X_n) alors π' est une distribution stationnaire pour la suite (Y_n) tel que

$$\pi' = (\pi'_y)_{y\in F} \quad \textit{et} \quad \pi'_y = \sum_{f(j)=y} \pi_j.$$

Preuve. *$\pi' = (\pi'_y)_{y\in F}$ est une distribution stationnaire si et seulement si : $\sum_{y\in F} \pi'_y = 1$ et $\sum_{x\in F} \pi'_x \cdot q_{x,y} = \pi'_y$.*

$\triangleright \sum_{y\in F} \pi_y = \sum_{y\in F} \sum_{f(j)=y} \pi_{f^{-1}(y)} = \sum_{j\in E} \pi_j = 1 \quad car \quad E = \bigcup_{y\in F} f^{-1}(\{y\}).$

$\triangleright \sum_{x\in F} \pi'_x \cdot q_{x,y} = \sum_{i\in E} \sum_{f(j)=y} \pi_i p_{i,j} = \sum_{f(j)=y} \pi_j = \sum_{f(j)=y} \pi_{f^{-1}(y)} = \sum_{f(j)=y} \pi_y = \pi'_y$

Exemple : On suppose que $(X_n)_{n>0}$ est une chaîne de Markov sur $E = \{1,2,3\}$ de matrice de transition P :

$$P = \begin{pmatrix} 0 & 1-\alpha & \alpha \\ 1-\alpha & 0 & \alpha \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

de loi initiale $\mu_0 = (1/3, 1/3, 1/3)$, et de distribution stationnaire $\pi = (\frac{1}{2+3\alpha}, \frac{1}{2+3\alpha}, \frac{3\alpha}{2+3\alpha})$

Soit $f : E \to F = \{1,2\}$ donnée par : $f(1) = 1$ et $f(2) = f(3) = 2$ et $(Y_n)_n$ est une chaîne de Markov image par f.

$$\begin{aligned} q_{11} &= P(Y_{n+1} = 1/Y_n = 1) \\ &= P(X_{n+1} = 1/X_n = 1) \\ &= p_{11} = 0 \end{aligned}$$

$$\begin{aligned} q_{12} &= P(Y_{n+1} = 2/Y_n = 1) \\ &= P(X_{n+1} \in \{2,3\}/X_n = 1) \\ &= \frac{P(X_{n+1} \in \{2,3\}; X_n = 1)}{P(X_n = 1} \\ &= \frac{P(X_{n+1} = 2, X_n = 1) + P(X_{n+1} = 3, X_n = 1)}{P(X_n = 1)} \\ &= \frac{P(X_n = 1)P(X_{n+1} = 2/X_n = 1) + P(X_n = 1)P(X_{n+1} = 3/X_n = 1)}{P(X_n = 1)} \\ &= P(X_{n+1} = 2/X_n = 1) + P(X_{n+1} = 3/X_n = 1) \\ &= p_{12} + p_{13} = 1 \end{aligned}$$

$$\begin{aligned} q_{21} &= P(Y_{n+1} = 1/Y_n = 2) \\ &= P(X_{n+1} = 1/X_n \in \{2,3\}) \\ &= \frac{P(X_{n+1} = 1; X_n \in \{2,3\})}{P(X_n \in \{2,3\})} \\ &= \frac{P(X_{n+1} = 1; X_n = 2) + P(X_{n+1} = 1; X_n = 3)}{P(X_n = 2) + P(X_n = 3)} \\ &= \frac{P(X_n = 2)P(X_{n+1} = 1/X_n = 2) + P(X_n = 3)P(X_{n+1} = 1/X_n = 3)}{P(X_n = 2) + P(X_n = 3)} \\ &= \frac{P(X_n = 2)p_{21} + P(X_n = 3)p_{31}}{P(X_n = 2) + P(X_n = 3)} \\ &= \frac{1}{3\alpha + 1} \end{aligned}$$

$$\begin{aligned} q_{22} &= P(Y_{n+1} = 2/Y_n = 2) \\ &= P(X_{n+1} \in \{2,3\}/X_n \in \{2,3\} \\ &= \frac{P(X_{n+1} \in \{2,3\}; X_n \in \{2,3\}}{P(X_n \in \{2,3\})} \\ &= \frac{P(X_n = 2)P(X_{n+1} = 2/X_n = 2) + P(X_n = 3)P(X_{n+1} = 2/X_n = 3)}{P(X_n = 2) + P(X_n = 3)} \\ &+ \frac{P(X_n = 2)P(X_{n+1} = 3/X_n = 2) + P(X_n = 3)P(X_{n+1} = 3/X_n = 3)}{P(X_n = 2) + P(X_n = 3)} \\ &= \frac{3\alpha}{1 + 3\alpha} \end{aligned}$$

Alors l'image de P est Q avec :

$$Q = \begin{pmatrix} 0 & 1 \\ 1/3\alpha + 1 & 3\alpha/3\alpha + 1 \end{pmatrix}$$

La distribution stationnaire de Y_n est :

$\pi' = (\pi'_y)_{y\in F}$ tel que $\pi'_y = \sum\limits_{f(j)=y} \pi_j$

$\pi'_1 = \sum\limits_{f(j)=1} \pi_j = \pi_1 = \dfrac{1}{2+3\alpha}$

$\pi'_2 = \sum\limits_{f(j)=2} \pi_j = \pi_2 + \pi_3 = \dfrac{1+3\alpha}{2+3\alpha}$

Donc

$$\pi' = (\frac{1}{2+3\alpha}, \frac{1+3\alpha}{2+3\alpha})$$

2.4.1 Caractérisation d'une chaîne de Markov image

Proposition 2.12.

Si $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène irréductible vérifie la condition de la proposition et f une application surjective tel que $Y_n = f(X_n)$ alors la chaîne de Markov $(f(X_n))_{n\geq 0}$ est aussi irréductible.

Preuve. *Soient $x, y \in F$, $i, j \in E$ tel que $f(i) = x$ et $f(j) = y$.*
$(X_n)_n$ une chaîne de Markov irréductible $\Longleftrightarrow \forall\, i, j \in E, \exists\, n, m \in \mathbb{N}^*, p^{(n)}_{i,j} > 0$ et $p^{(m)}_{j,i} > 0$;
$(Y_n)_n$ une chaîne de Markov irréductible $\Longleftrightarrow \forall\, x, y \in F, \exists\, n, m \in \mathbb{N}^*, q^{(n)}_{x,y} > 0$ et $q^{(m)}_{y,x} > 0$.
On a :

$$q^{(n)}_{x,y} = \sum_{j\in E/f(j)=y} p^{(n)}_{i,j} > 0 \quad \textit{(Car } (X_n) \textit{ irréductible).}$$

$$q^{(m)}_{x,y} = \sum_{j\in E/f(j)=y} p^{(m)}_{i,j} > 0 \quad \textit{(Car } (X_n) \textit{ irréductible).}$$

Donc :$(Y_n)_n$ est irréductible.

Proposition 2.13.

Soient $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène et f une application surjective tel que $Y_n = f(X_n)$ une chaîne de Markov.
Soient $x \in F$ et $i \in E$ tel que $i \in f^{-1}(\{x\})$ alors :
1) i état récurrent $\Rightarrow$ x récurrent.
2) i état transitoire $\Rightarrow$ x transitoire.

Preuve. *Soit $x \in F$, alors $f^{-1}(\{x\}) \neq \emptyset \Rightarrow \exists\, i \in f^{-1}(\{x\}))$.*
1) Si i état récurrent alors

$$\sum_{n=1}^{+\infty} p^{(n)}_{i,i} = +\infty,$$

donc

$$\sum_{n=1}^{+\infty} q^{(n)}_{x,x} = \sum_{n=1}^{+\infty} \sum_{j\in E/f(j)=y} p^{(n)}_{j,j} = \sum_{n=1}^{+\infty} p^{(n)}_{j,j} = +\infty$$

(car i état récurrente) alors x est récurrente.
2) Si i état transitoire alors $\sum_{n=1}^{+\infty} p_{j,j}^{(n)} < +\infty$*, donc*

$$\sum_{n=1}^{+\infty} q_{x,x}^{(n)} = \sum_{n=1}^{+\infty} \sum_{j \in E/f(j)=y} p_{j,j}^{(n)} = \sum_{n=1}^{+\infty} p_{j,j}^{(n)} < +\infty$$

(car i état transitoire) alors x est transitoire.

2.4.2 Le comportement asymptotique d'une chaîne de Markov image

Soient $(X_n)_n$ une chaîne de Markov d'espace d'états E fini, de distribution limite π et f une application surjective tel que $f : E \to F$ et $Y_n = f(X_n)$.
Si la chaîne $(Y_n)_n$ satisfait la condition nécessaire et suffisante alors $(Y_n)_n$ est une chaîne de Markov image d'espace d'états F, de distribution limite $\pi' = \sum_{f(j)=y} \pi_j$ et de matrice de transition $Q = q_{x,y} = \sum_{j \in E/f(j)=y} p_{i,j}$.

2.4.2.1 Chaîne de Markov irréductible apériodique

Proposition 2.14.

Si $(X_n)_n$ est irréductible apériodique alors $(Y_n)_n$ est une chaîne de Markov irréductible apériodique et

$$\forall x, y \in F, \lim_{n\to+\infty} q_{x,y}^{(n)} = \lim_{n\to+\infty} \sum_{j \in E/f(j)=y} p_{i,j}^{(n)} = \pi_{f^{-1}(y)}$$

Si la chaîne $(Y_n)_n$ est ergodique on a :

$$\lim_{n\to+\infty} Q^{(n)} = Q^*.$$

Ou Q^* est une matrice dans tout les lignes égales à π'.

2.4.2.2 Chaîne de Markov irréductible périodique

Proposition 2.15.

Si $(X_n)_n$ est irréductible périodique alors $(Y_n)_n$ est une chaîne de Markov irréductible périodique de période $(d > 1)$ et

$$\lim_{n\to+\infty} r_{x,y}^{(n)} = \lim_{n\to+\infty} \sum_{j \in E/f(j)=y} p_{i,j}^{(nd)} = \begin{cases} d\pi'_j & \text{si } f^{-1}(x), f^{-1}(y) \text{appartient à même classe ;} \\ 0 & \text{si non.} \end{cases}$$

$$r_{x,y}^{(n)} = \sum_{j \in E/f(j)=y} (p_{i,j}^{(d)})^n = \sum_{j \in E/f(j)=y} p_{i,j}^{(nd)}.$$

2.4.2.3 Chaîne de Markov réductible

Théorème 2.1.

Si $(Y_n)_n$ une chaîne de Markov réductible ne contient pas une seule classe d'équivalence. Soit un état $x \in F$,la probabilité que $(Y_n)_n$ se trouve dans un état y à l'étape n est :

$$\lim q_{x,y}^{(n)} = \begin{cases} 0 & \textit{Si } f^{-1}(y) \textit{ est transitoire.} \\ 0 & \textit{Si } f^{-1}(x) \textit{ et } f^{-1}(y) \textit{ appartient à deux classes différents.} \\ \pi_f^{-1}(y) & \textit{Si } f^{-1}(x) \textit{ et } f^{-1}(y) \textit{ récurrentes appartient à même classe.} \\ \pi_{f^{-1}(y)} \cdot A_{xc} & \textit{Si } f^{-1}(x) \textit{ est transitoire et } f^{-1}(y) \textit{ est récurrente.} \end{cases}$$

Où $\quad A_{xc} = \sum_{y \in C} p_{f^{-1}(x),f^{-1}(y)} + \sum_{k \in T} p_{f^{-1}(x),k} \cdot A_{kc} = \textit{Probabilité de } x \textit{ vers } y.$

2.5 Critères dérivées pour une chaîne de Markov image

Dans cette partie on va montrer que si un critère dérivée est applicable pour une chaîne de Markov, alors ce résultat reste vraie pour son image si cette image est une chaîne de Markov. Autrement dit, si il existe une fonction V qui vérifiée les conditions d'un critère pour une chaîne de Markov, alors il existe une fonction W qui vérifiée ces conditions pour la chaîne de Markov image.

On considère une chaîne de Markov image $Y_n = f(X_n)_{n\in\mathbb{N}}$ et $f : E \to F$ une fonction, et P un noyau de transition irréductible sur E.

2.5.1 Le cas d'une fonction injective(bijective)

Soient $V : E \to \mathbb{R}$ une fonction et $W : F \to \mathbb{R}$ avec $W(x) = V(f^{-1}(x)) = V \circ f^{-1}(x)$

Proposition 2.16.
$\forall x \in F$ et $i \in E, f(i) = x$:
(A) pour tout $x \in F$, $W(x) \geq 0$.
(B) W n'est pas constante sur F.
(C) $\mathbb{E}_x(W(Y_1)) = \mathbb{E}_i(V(X_1))$.
(D) $\mathbb{E}_x(W(Y_0)) = V(i) = \mathbb{E}_i(V(X_0))$.
(E) $\mathbb{E}_x(W(Y_1) - W(Y_0)) = \mathbb{E}_i(V(X_1) - V(X_0))$.

Preuve.
(A) on a $V(i) \geq 0$ alors $W(x) \geq 0$.
(B) comme V n'est pas constante sur E alors W n'est pas constante sur F.
(C) pour tout $x, y \in F$ et $i, j \in E$:

$$\begin{aligned}\mathbb{E}_x(W(Y_1)) &= \sum_{y\in F} W(y)Q_{x,y}\\ &= \sum_{y\in F/f(y)=j} V \circ f^{-1}(y)P_{f^{-1}(x)f^{-1}(y)}\\ &= \sum_{j\in E} V(j)P_{i,j}\\ &= \mathbb{E}_i(V(X_1)).\end{aligned}$$

(D) pour tout $x, y \in F$ et $i \in E$ tel que $f(i) = x$:

$$\begin{aligned}\mathbb{E}_x(W(Y_0)) &= \sum_{y\in F} W(y)P(Y_0 = y/Y_0 = x)\\ &= \sum_{y\in F/f(j)=y} V \circ f^{-1}(y)P(X_0 = f^{-1}(y)/X_0 = f^{-1}(x))\\ &= \sum_{j\in E} V(j)P(X_0 = j/X_0 = i)\\ &= V(i) = \mathbb{E}_i(V(X_0)).\end{aligned}$$

(E) pour tout $x \in F$ et $i \in E$:

$$\begin{aligned}\mathbb{E}_x(W(Y_1) - W(Y_0)) &= \mathbb{E}_x(W(Y_1)) - \mathbb{E}_x(W(Y_0)) \\ &= \mathbb{E}_i(V(X_1)) - V(i) \\ &= \mathbb{E}_i(V(X_1)) - \mathbb{E}_i(V(X_0)) \\ &= \mathbb{E}_i(V(X_1) - V(X_0)).\end{aligned}$$

2.5.1.1 Critère de non-récurrence positive

Proposition 2.17. *supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) V n'est pas constante sur E ;
(3) pour tout $i \in E$, $\mathbb{E}_i(V(X1)) < +\infty$;
(4) pour tout $i \in E$, $\mathbb{E}_i(V(X_1) - V(X_0)) \geq 0$;
(5) il existe $c \geq 0$ tel que, pour tout $i \in E$, $\mathbb{E}_i \mid V(X_1) - V(X_0) \mid \leq c$.
Alors la chaîne image Y_n est soit transiente, soit récurrente nulle.

2.5.1.2 Critère de transience

Proposition 2.18. *Supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) V n'est pas constante sur E ;
(3) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$;
(4) pour tout $i \in E$, $\mathbb{E}_i(V(X_1) - V(X_0)) \leq 0$;
Alors la chaîne image Y_n est soit transiente. La réciproque est vraie.

2.5.1.3 Critère de récurrence

Proposition 2.19. *Supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$;
(3) il existe un sous-ensemble fini C de E, tel que, pour tout $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq 0;$$

(4) pour tout K, l'ensemble $\{i \in E, V(i) \geq K\}$ est fini.
Alors la chaîne image Y_n est récurrente.

2.5.1.4 Critère de récurrence positive

Proposition 2.20. *Supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*

(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$;
(3) il existe $\varepsilon > 0$ et un sous-ensemble fini C de E, tel que, pour tout $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq -\varepsilon;$$

Alors la chaîne image Y_n est récurrente positive. La réciproque est vraie.

2.5.1.5 Critère de d'ergodicité géométrique

Proposition 2.21. *Supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$;
(3) il existe $\varepsilon > 0$ et un sous-ensemble fini C de E, tel que, pour tout $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq -\varepsilon V(i);$$

Alors il existe $i \in E$, $a, b > 0$ tels que, tout $n \geq 0$,

$$P_i(T_1(i) \geq n) \leq a \cdot exp(-bn).$$

Avec $T_1(i)$ est un temps d'arrêt.

Preuve. *Comme la fonction $W(x) = V \circ f^{-1}(x) = V(i)$ et d'après la proposition 2.16 alors tout les propriétés des critères dérivées pour une chaîne de Markov sont applicables sur la chaîne de Markov image.*

2.5.2 Le cas d'une fonction surjective

Dans cette partie, nous allons aborder le cas d'une application " non injective ", et donc "non bijective".

2.5.2.1 Critère de non-récurrence positive

Proposition 2.22. *Supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) V n'est pas constante sur E ;
(3) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$;
(4) pour tout $i \in F$, $\mathbb{E}_i(V(X_1) - V(X_0)) \geq 0$;
(5) il existe $c \geq 0$ tel que, pour tout $i \in E$, $\mathbb{E}_i \mid V(X_1) - V(X_0) \mid \leq c$.
Par conséquent il existe une fonction $W : F \to \mathbb{R}$ vérifiant les propriétés suivantes :
(1) pour tout $x \in F$, $W(x) \geq 0$;
(2) W n'est pas constante sur F ;
(3) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1)) < +\infty$;

(4) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1) - W(Y_0)) \geq 0$;
(5) il existe $c \geq 0$ tel que, pour tout $x \in F$, $\mathbb{E}_x \mid W(Y_1) - W(Y_0) \mid \leq c$.
Alors la chaîne image Y_n est soit transiente, soit récurrente nulle.

Preuve. *Pour $y \in F$, on pose*

$$W(y) = \sum_{f(j)=y} V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}}.$$

Montrons que $\forall\, y \in F, \exists\, M, m \in E : V(m) \leq W(y) \leq V(M)$.
$\triangleright$ *Soit $M \in E$ tel que : $V(M) = \max_{f(j)=y} V(j)$*
Pour $j \in E$ et $f(j) = y$:

$$V(j) \leq V(M) \Rightarrow V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \leq V(M) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}}$$

$$\Rightarrow W(y) \leq V(M) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}}.$$

On pose : $\max_{k} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} = \frac{P_{M',j}}{\sum_{f(l)=y} P_{M',l}}$

$$W(y) \leq V(M) \frac{\sum_{f(j)=y} P_{M',j}}{\sum_{f(l)=y} P_{M',l}} = V(M),$$

Par conséquent :

$$W(y) \leq V(M).$$

$\triangleright$ *Soit $m \in E$ tel que : $V(m) = \min_{f(j)=y} V(j)$*
Pour $j \in E$ et $f(j) = y$:

$$V(j) \geq V(m) \Rightarrow V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \geq V(m) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}}$$

$$\Rightarrow W(y) \geq V(m) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}}$$

$$W(y) \geq V(m) \frac{\sum_{f(j)=y} P_{M',j}}{\sum_{f(l)=y} P_{M',l}} = V(m),$$

Par conséquent :

$$W(y) \geq V(m).$$

(1) $V \geq 0$ *donc* $W \geq 0$.
(2) V *n'est pas constante sue* E *alors* W *n'est pas constante sur* F.
(3) pour tout $x \in F$ *tel que* $f(i) = x$, $\quad \mathbb{E}_x(W(Y_1)) < +\infty$
On pose que

$$\sum_y \sum_{f(j)=y} V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \sum_{f(h)=y} P_{i,k} < +\infty.$$

Alors

$$\mathbb{E}_x(W(Y_1)) = \sum_y W(y) Q_{x,y} = \sum_y \sum_{f(j)=y} V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \sum_{f(h)=y} P_{i,k} < +\infty.$$

(4)Pour tout $i \in E$ *et* $x \in F$ *:*

$$\begin{aligned}
\mathbb{E}_x(W(Y_1)) &= \sum_y W(y) Q_{x,y} \\
&= \sum_y \sum_{f(j)=y} V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \sum_{f(h)=y} P_{i,k} \\
&\geq \sum_y \sum_{f(j)=y} V(j) P_{i,j} \\
&= \sum_{j \in E} V(j) P_{i,j} = \mathbb{E}_i(V(X_1)).
\end{aligned}$$

Donc

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) \geq \mathbb{E}_{i_{Mx}}(V(X_1) - V(X_0)).$$

D'où le résultat : Si $\forall\, i \in E$

$$\mathbb{E}_i(V(X_1) - V(X_0)) \geq 0.$$

Alors $\forall\, x \in F$

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) \geq 0.$$

(5) $\mathbb{E}_x \mid W(Y_1) - W(Y_0) \mid \leq \mathbb{E}_x \mid W(Y_1) \mid - V(i_{mx})$ *avec* $V(i_{mx}) = \min_{f(i)=x} V(i)$

$$\mathbb{E}_x \mid W(Y_1) \mid - V(i_{mx}) = \sum_y \sum_{f(j)=y} V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \sum_{f(h)=y} P_{i,k}$$

$$\begin{aligned}
\Longleftrightarrow\ \mathbb{E}_x \mid W(Y_1) \mid &\leq \sum_y \sum_{f(j)=y} V(j) \max_{k \in E} \frac{P_{k,j}}{\sum_{f(l)=y} P_{k,l}} \sum_{f(h)=y} P_{i,k} - V(i_{mx}) \\
&\leq \sum_j V(j) P_{i_{xm}} - V(i_{mx}) \\
&= \mathbb{E}_{i_{xm}} \mid V(X_1) - V(X_0) \mid \leq c.
\end{aligned}$$

Donc : $\mathbb{E}_x \mid W(Y_1) - W(Y_0) \mid \leq c$.

2.5.2.2 Critère de transience

Proposition 2.23. *Supposons qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$;
(2) V n'est pas constante sur E ;
(3) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$;
(4) pour tout $i \in E$, $\mathbb{E}_i(V(X_1) - V(X_0)) \leq 0$;
Par conséquent il existe une fonction $W : F \to \mathbb{R}$ vérifiant les propriétés suivantes :
(1) pour tout $x \in F$, $W(x) \geq 0$;
(2) W n'est pas constante sur F ;
(3) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1)) < +\infty$;
(4) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1) - W(Y_0)) \leq 0$;
Alors la chaîne image Y_n est soit transiente. La réciproque est vraie.

Preuve. *Supposons qu'il existe une fonction $W : E \to \mathbb{R}$ avec*

$$W(y) = \begin{cases} V(j) & si \quad A_y = \{j\}; \\ \sum\limits_{j \in A_y} V(j) \min\limits_{i \in E} \dfrac{P_{i,j}}{\sum\limits_{k \in A_y} P_{i,j}} & si\ Card A_y > 1. \end{cases}$$

avec $A_y = \{j \in E, f(j) = y\}$.
(1) pour tout $x \in F$, $W(x) \geq 0$ car $V(j) \geq 0$ et $P_{i,j} \geq 0$.
(2) W n'est pas constante sur F car $V(j)$ n'est pas constante.
(3) pour tout $x \in F$,

$$\begin{aligned}
\mathbb{E}_x(W(Y_1)) &= \sum_{y \in F} W(y) Q_{x,y} \\
&= \sum_{y \in F(A_y = \{j\})} W(y) Q_{x,y} + \sum_{Card A_y > 1} W(y) Q_{x,y} \\
&= \sum_{f(j) = y(Card A_y = 1)} V(j) P_{i,j} + \sum_{Card A_y > 1} \sum_{j \in A_y} V(j) \min_{i \in E} \frac{P_{i,j}}{\sum\limits_{k \in A_y} P_{i,k}} \sum_{m \in A_y} P_{i,m} \\
&\leq \sum_{f(j) = y(Card A_y = 1)} V(j) P_{i,j} + \sum_{Card A_y > 1} \sum_{j \in A_y} V(j) \frac{P_{i,j}}{\sum\limits_{k \in A_y} P_{i,k}} \sum_{m \in A_y} P_{i,m} = \sum_E V(j) P_{i,j} \\
&= \mathbb{E}_x(V(X_1)) < +\infty.
\end{aligned}$$

Donc $\mathbb{E}_x(W(Y_1)) < +\infty$.
(4) pour tout $x \in F$, et $\mathbb{E}_x(W(Y_0)) = \sum\limits_{y \in F} W(y) P(Y_0 = 0 / Y_0 = x) = W(x)$.

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) = \mathbb{E}_x(W(Y_1)) - \mathbb{E}_x(W(Y_0)) \leq \mathbb{E}_i(V(X_1)) - W(x) \leq 0.$$

2.5.2.3 Critère de récurrence

Proposition 2.24. *Supposons que $\forall\, x \in F$, A_x est fini, et qu'il existe une fonction $V : E \to \mathbb{R}$ vérifiant les propriétés suivantes :*

(1) pour tout $i \in E$, $V(i) \geq 0$*;*
(2) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$*;*
(3) il existe un sous-ensemble fini C *de* E*, tel que, pour tout* $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq 0;$$

(4) pour tout K*, l'ensemble* $\{i \in E, V(i) \geq K\}$ *est fini.*
Par conséquent il existe une fonction $W : F \to \mathbb{R}$ *vérifiant les propriétés suivantes :*
(1) pour tout $x \in F$, $W(x) \geq 0$*;*
(2) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1)) < +\infty$*;*
(3) il existe un sous-ensemble fini D *de* F*, tel que, pour tout* $x \notin D$,

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) \leq 0;$$

(4) pour tout K*, l'ensemble* $\{x \in F, W(x) \geq K\}$ *est fini.*
Alors la chaîne image Y_n *est récurrente.*

Preuve. *On pose* $W(x) = \sum_{f(j)=x} V(j) \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}}$.
Pour les propriétés (1),(2) et (3) est le même démonstration comme le proposition précédente.
(4) pour tout K*, on pose* $\alpha = \min_{x/f(j)=x} \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}}$ *et* $\alpha' = \min_x CardA_x$

On a : pour tout $j \in E, V(j) \geq K$
$\Rightarrow V(j) \cdot \alpha_x \geq K \cdot \alpha_x$
$\Rightarrow V(j) \cdot \alpha \geq K \cdot \alpha$
$\Rightarrow \sum_{f(j)=x} V(j)\alpha \geq \sum_{f(j)=x} K \cdot \alpha = K \cdot \alpha \cdot CardA_x$
$\Rightarrow W(x) \geq K \cdot \alpha \cdot \alpha'$

Donc : $\forall K, \exists K' = \frac{K}{\alpha \cdot \alpha}$ $\{x \in F, W(x) \geq K\}$ *est fini.*

2.5.2.4 Critère de récurrence positive

Proposition 2.25. *Supposons qu'il existe une fonction* $V : E \to \mathbb{R}$ *vérifiant les propriétés suivantes :*
(1) pour tout $i \in E$, $V(i) \geq 0$*;*
(2) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$*;*
(3) il existe $\varepsilon > 0$ *et un sous-ensemble fini* C *de* E*, tel que, pour tout* $i \notin C$,

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq -\varepsilon;$$

Par conséquent il existe une fonction $W : F \to \mathbb{R}$ *vérifiant les propriétés suivantes :*
(1) pour tout $x \in F$, $W(x) \geq 0$*;*
(2) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1)) < +\infty$*;*
(3) il existe $\varepsilon > 0$ *et un sous-ensemble fini* D *de* F*, tel que, pour tout* $x \notin D$,

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) \leq -\varepsilon;$$

Alors la chaîne image Y_n est récurrente positive. La réciproque est vraie.

Preuve. *On pose* $W(x) = \sum_{f(j)=x} V(j) \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}}$,

Alors pour (1) et (2) sont déjà vérifiées dans la proposition 2.23

(3) il existe $\varepsilon > 0$ *et un sous-ensemble fini* $D = f(C)$ *de* F*, tel que, pour tout* $x \notin D$*,*

$$\begin{aligned}\mathbb{E}_x(W(Y_1) - W(Y_0)) &= \mathbb{E}_x(W(Y_1)) - \mathbb{E}_x(W(Y_0)) \\ &= \mathbb{E}_x(W(Y_1)) - W(x).\end{aligned}$$

On pose $V(j_{Mx}) = \max_{f(j)=x} V(j)$ *et* $V(j_{mx}) = \min_{f(j)=x} V(j)$.

$$\forall j, f(j) = x: \quad V(j_{mx}) \leq V(j) \leq V(j_{Mx})$$

$$\Rightarrow V(j_{mx}) \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}} \leq V(j) \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}} \leq V(j_{Mx}) \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}}.$$

$$\Rightarrow V(j_{mx}) \sum_{f(j)=x} \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}} \leq W(x) \leq V(j_{Mx}) \sum_{f(j)=x} \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}}.$$

Alors

$$\begin{aligned}0 < V(j_{mx}) \leq W(x) \leq V(j_{Mx}) \\ \Rightarrow -V(j_{Mx}) \leq -W(x) \leq -V(j_{mx}).\end{aligned}$$

Donc,

$$\mathbb{E}_x(W(Y_1)) - W(x) \leq \mathbb{E}_{j_{mx}}(V(X_1)) - V(j_{mx}) = \mathbb{E}_{j_{mx}}(V(X_1)) - V(X_0)) < -\varepsilon$$

$$\Rightarrow \mathbb{E}_x(W(Y_1)) - W(x) \leq -\varepsilon.$$

2.5.2.5 Critère de d'ergodicité géométrique

Proposition 2.26. *Supposons qu'il existe une fonction* $V : E \to \mathbb{R}$ *vérifiant les propriétés suivantes :*

(1) pour tout $i \in E$, $V(i) \geq 0$*;*

(2) pour tout $i \in E$, $\mathbb{E}_i(V(X_1)) < +\infty$*;*

(3) il existe $\varepsilon > 0$ *et un sous-ensemble fini* C *de* E*, tel que, pour tout* $i \notin C$*,*

$$\mathbb{E}_i(V(X_1) - V(X_0)) \leq -\varepsilon V(i);$$

Alors il existe $i \in E$*,* $a, b > 0$ *tels que, tout* $n \geq 0$*,*

$$P_i(T_1(i) \geq n) \leq a \cdot \exp(-bn).$$

Avec $T_1(i)$ *est un temps d'arrêt. Par conséquent il existe une fonction* $W : F \to \mathbb{R}$ *vérifiant les propriétés suivantes :*

(1) pour tout $x \in F$, $W(x) \geq 0$;
(2) pour tout $x \in F$, $\mathbb{E}_x(W(Y_1)) < +\infty$;
(3) il existe $\varepsilon > 0$ et un sous-ensemble fini D de F, tel que, pour tout $x \notin D$,

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) \leq -\varepsilon W(x);$$

Alors il existe $x \in F$, $a, b > 0$ tels que, tout $n \geq 0$,

$$P_x(T_1'(x) \geq n) \leq a \cdot exp(-bn).$$

Avec $T_1'(x)$ est un temps d'arrêt.

Preuve. *En utilisant la même fonction définie* $W(x) = \sum_{f(j)=x} V(j) \min_k \frac{P_{k,j}}{\sum_{f(l)=x} P_{k,l}}$.

Pour (1) et (2) sont déjà vérifiées dans la proposition 2.23
(3) pour tout $i \notin C \iff f(i) = x \notin D = f(C)$
f surjective et C fini donc D fini
Soit $x \notin D$,

$$\forall j \in E, f(j) = x \Rightarrow j \notin C$$

On a

$$0 \leq V(j_{mx}) \leq W(x) \leq V(j_{Mx})$$

Donc

$$\varepsilon V(j_{mx}) \leq \varepsilon W(x) \leq \varepsilon V(j_{Mx})$$

Et

$$-\varepsilon V(j_{Mx}) \leq -\varepsilon W(x) \leq -\varepsilon V(j_{mx})$$

Pour $j_{Mx} \notin C$ et d'après la condition (3) de cette proposition on a :

$$\mathbb{E}_{j_{Mx}}(V(X_1) - V(X_0)) \leq -\varepsilon V(j_{Mx})$$

et

$$\mathbb{E}_x(W(Y_1)) - W(Y_0)) \leq \mathbb{E}_{j_{Mx}}(V(X_1) - V(X_0)) \leq -\varepsilon V(j_{Mx}).$$
$$-\varepsilon V(j_{Mx}) \leq -\varepsilon W(x).$$

Donc

$$\mathbb{E}_x(W(Y_1) - W(Y_0)) \leq -\varepsilon W(x).$$

D'où le résultat.

Conclusion :

On termine ce chapitre par la conclusion suivante :

• Si f est injective, alors $f(X_n)$ est une chaîne de Markov.

• Si f est surjective, $f(X_n)$ est une chaîne de Markov si et ssi $\forall A \subset F$:

$$\forall\, i,k \in E, f(i) = k : \sum_{j\in A} p_{i,j} = \sum_{j\in A} p_{k,j}.$$

• Si f n'est pas injective et n'est pas surjective, il suffit de prendre Imf au lieu de F pour rendre l'application surjective dans ce cas $f(X_n)$ est une chaîne de Markov d'espace d'états Imf, avec la condition sur les probabilités de transition.

• La chaîne de Markov $(f(X_n))_n$ posséde les mêmes propriétés que la chaîne de Markov $(X_n)_n$ (la récurrence,la transience,la comportement asymptotique...).

• Si un critère dérivée est vraie pour X_n, alors ce résultat reste vraie pour la chaîne de Markov image Y_n.

Chapitre 3

Application en Matlab

L'objectif de ce chapitre est d'illustrer numériquement les notions des chaînes de Markov selon un logiciel très connu dans l'espace des calculs rapides . Avant de passer à la simulation dans MATLAB , il faut présenter l'intérêt de cet outil de programmation calcul numérique . Alors la question qui se pose ici est de savoir pourquoi nous choisissons Matlab ?

3.1 Introduction aux Matlab

MATLAB est une abréviation de **MATrix LABoratory**, était destiné à faciliter l'accès au logiciel matriciel. MATLAB est un langage de calcul scientifique de haut niveau et un environnement interactif pour le développement d'algorithmes, la visualisation et l'analyse de données, ou encore le calcul numérique.

3.1.1 Caractéristiques Matlab

• Résoudre des problèmes de calcul scientifique plus rapidement.

• MATLAB offre un certain nombre de fonctionnalités pour la documentation et le partage de votre travail.

• Des boîtes à outils supplémentaires (collections de fonctions MATLAB à vocation spécifique, disponibles séparément) élargissent l'environnement MATLAB pour résoudre des catégories particulières de problèmes dans ces domaines d'applications : SIMULINK, réseaux de neurones et la logique floue.

3.1.2 Utilisation Matlab

MATLAB utilisé dans une grande variété d'applications :

• Le traitement du signal et d'images.

• Les communications.

• La conception de systèmes de contrôle.

• Les tests et les mesures.

• La modélisation et l'analyse financière.

3.1.3 Principales fonctionnalités

• Langage de haut niveau pour le calcul scientifique.

• Environnement de développement pour la gestion du code, des fichiers et des données.

• Outils interactifs pour l'exploration itérative, la conception et la résolution de problèmes.

• Fonctions mathématiques pour l'algèbre linéaire, les statistiques, l'analyse de Fourier, le filtrage, l'optimisation et l'intégration numérique.

• Fonctions graphiques 2D et 3D pour la visualisation des données.

• Outils pour la construction d'interfaces graphiques personnalisées.

• Fonctions pour l'intégration d'algorithmes développés en langage MATLAB, dans des applications et langages externes (C/C++, Fortran, Java, COM et Microsoft Excel).

3.2 Un exemple pratique

[**Marche aléatoire sur** $\mathbb{Z}/3\mathbb{Z}$]

Un joueur fréquente 3 salles des jeux numérotés 1, 2 et 3. Chaque jour il choisit l'un des deux salles où il n'est pas allé la veille suivant une même probabilité 1/2. Le premier jour, jour 0, il choisit l'un des trois salles des jeux selon une loi $\mu_0 = (1/2, 0, 1/2)$ sur $E = \{1,2,3\}$. On note X_n la variable aléatoire égale au numéro du salle des jeux fréquenté par le joueur le jour n. On admet que $(X_n)_{n\geq 0}$ est une chaîne de Markov de loi initiale μ et de distribution stationnaire π.

La matrice de transition

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

Dans ce chapitre nous utiliserons toujours la matrice mentionnée dans l'exemple pratique.

Remarque : On note $\mu = Mu$ dans l'algorithme que nous avons les traiter.

L'algorithme :

```
clc
clear all
% Définir la matrice de transition d'une chaine de markov
n=input('Donnes la taille=');
for i=1:n
    for j=1:n
        P(i,j)=input ('Donner les valeur de la matrice P:');
    end
end
P

% Donner la loi initiale de cette chaine
 for i=1:n
        Mu(i)=input ('Donner le vecteur de la loi initiale Mu:');
 end
 Mu
```

FIGURE 3.1 – Définir la chaîne de Markov X_n

Exécution :

```
Donnes la taille=
3
Donner les valeur de la matrice P:
0
Donner les valeur de la matrice P:
1/2
Donner les valeur de la matrice P:
1/2
Donner les valeur de la matrice P:
1/2
Donner les valeur de la matrice P:
0
Donner les valeur de la matrice P:
1/2
Donner les valeur de la matrice P:
1/2
Donner les valeur de la matrice P:
1/2
Donner les valeur de la matrice P:
0
```

```
P =

         0    0.5000    0.5000
    0.5000         0    0.5000
    0.5000    0.5000         0

Donner le vecteur de la loi initiale Mu:
1/2
Donner le vecteur de la loi initiale Mu:
0
Donner le vecteur de la loi initiale Mu:
1/2

Mu =

    0.5000         0    0.5000
```

FIGURE 3.2 – Exécution de la chaîne de Markov

3.2.1 Le calcule de la distribution stationnaire de la chaîne de Markov

Vérification de l'existence dune distribution stationnaire

L'algorithme :

```
clc
clear all
P=[0,1/2,1/2;1/2,0,1/2;1/2,1/2,0];
Mu=[1/2,0,1/2];
% Vérivication de l'existance d' une distribution stationnaire
for k=1:100
    Pk=P^k;
    if Pk>0
        k %la valeur de la puissance laqualle Pk>0
        break
    end
end
Pk
if Pk>0
    disp('la chaine possède une DS ')
    N=1;
else
    disp('la chaine ne possède pas une DS')
end
```

FIGURE 3.3 – Vérification de l'existence dune distribution stationnaire de la chaîne X_n

Exécution :

```
k =

     2


Pk =

    0.5000    0.2500    0.2500
    0.2500    0.5000    0.2500
    0.2500    0.2500    0.5000

la chaine possède une DS
```

FIGURE 3.4 – Exécution de vérification de l'existence dune distribution stationnaire de la chaîne X_n

3.2.2 Caractérisation d'une chaîne de Markov

3.2.2.1 L'irréductibilité

L'algorithme :

```
P=[0,0.5,0.5;0.5,0,0.5;0.5,0.5,0];
P
Z=zeros(length(P));
Z
for i=1:length(P)
    for j=1:length(P)
        if P(i,j)>0 z(i,j)=1;
        end
    end
end
I=(eye(length(P))+z)^(length(P)-1);
I
if I>0
    disp('la chaîne de Markov est irréductible');
else
    disp('la chaîne de Markov est réductible ');
end
```

FIGURE 3.5 – Irréductibilité de la chaîne X_n

Exécution

```
P =

         0    0.5000    0.5000
    0.5000         0    0.5000
    0.5000    0.5000         0


Z =

     0     0     0
     0     0     0
     0     0     0


I =

     3     3     3
     3     3     3
     3     3     3

la chaîne de Markov est irréductible
```

FIGURE 3.6 – Exécution de L'irréductibilité de la chaîne X_n

3.2.2.2 Périodicité

L'algorithme :

```
P=[0,0.5,0.5;0.5,0,0.5;0.5,0.5,0]
n=10;
if ~isempty(find(diag(P)))
    disp('la chaîne de Markov est apériodique');
else
    I=P;k=0;
    for i=2:n
        I=I*P;
        if ~isempty(find(diag(I)))
            k=k+1;
            MP(k,1:length(P))=zeros(1,length(P));
            MP(k,find(diag(I))>0)=i;
        end

    end
    I
    MP
    d=0;
   for i=1:n
       d= min(MP,[],'all');
   end
   d
   disp('la chaîne de Markov est périodique');
end
% abs (eig(P))
```

FIGURE 3.7 – Périodicité de la chaîne X_n

Exécution :

```
I =

    0.3340    0.3330    0.3330
    0.3330    0.3340    0.3330
    0.3330    0.3330    0.3340


d =

     2

la chaîne de Markov est périodique
```

FIGURE 3.8 – Exécution de la périodicité de la chaîne X_n

3.2.3 Chaîne de Markov image

3.2.3.1 Définir la fonction *f*

Afin de pouvoir calculer l'image d'une chaîne de Markov. nous avons besoin de définir une fonction $f : E \to F$

L'algorithme :

```
clc
clear all
P=[0,1/2,1/2;1/2,0,1/2;1/2,1/2,0];
Mu=[1/2,0,1/2];
% Definir la fonction f de l'image de la chaîne de Markov
E=input ('Donnez l ensemble de depart E:')
F=input('Donnez l ensemble d arrive F:')
e=length(E);% calcule la dimension de l'ensemble E,
            % qui est egale au nombre de ses éléments
f=length(F);% calcule la dimension de l'ensemble F,
            % qui est egale au nombre de ses éléments
% determination de la nature de la fonction f
c=0;
if e <= f
    c=1
    disp('f est une fonction injective (bijective),Yn est une chaîne de Markov image ')
else
     c=2
     disp('f est une fonction surjective, il faut vérifier la CNS')
end
% initialisation des valeurs de f
X=input('Donnez l ensemble des antécédents Xi:')
Y=input('Donnez l ensemble des images Yi associe aux antécédent Xi :')
```

FIGURE 3.9 – Définition de la fonction f pour une chaîne de Markov image

Si la fonction f est surjective il faut vérifier les CNS

```
% vérification du caractère Markovien de Yn dans le cas surjective.
if c==2
   P1=0;
   P2=0;
   H=0;
   for i=1:e
       if Y(i)= Y(i+1)
           1=i;
           c=i+1;
       end
   end
   for i=1:e
       P1=P1+P(1,i); %calcul du premier terme.
       P2=P2+P(c,i); %calcul du second terme.
   end
   if P1==P2
       disp('Yn est une chaîne de Markov image.')
       H=1;
   else disp('Yn n est pas une chaîne de Markov image.')
       return
   end
end
```

FIGURE 3.10 – les CNS pour le cas surjective

3.2.3.2 Matrice de transition et distribution stationnaire de chaîne de Markov image

Soit $Y_n = f(X_n)$ est une chaîne de Markov, Nous calculons sa matrice de transition et sa distribution stationnaire.

Voici l'algorithme qui permet de la calculer quelque soit le cas dans lequel on se trouve :

L'algorithme :

```
%calculer l image de la chaine
dsQ=0;
ds=0;
if c==1 % si f est injective
    for i=1:e
        for j=1:f
            Q(i,j)=P(Y(i),Y(j))
        end
    end
    Q
    for i=1:f
        dsQ(i)=ds(Y(i))
    end
    dsQ
else if H==1 %si f est surjective
        for i=1:f
            for j=1:f
                if P(i,j)==0
                    Q(i,j)=P(i+1,j)
                else
                    Q(i,j)=P(i,j)
                end
            end
        end
        Q
        for i=1:l
            dsQ=[ds(i),dsQ];
        end
```

```
% calculer la distribution stationnaire image
verif =dsQ*Q
if verif== dsQ
    verif=1;
end
ver=0;
if verif>=0;
    ver=1;
end
q=0
for i=1:f
    q=q+dsQ(i)
end
if q==1
    if ver==1
        if verif==1
            disp('la distribution image est stationnaire ')
        else
            disp('la distribution image n est pas stationnaire')
end
end
end
```

FIGURE 3.11 – Matrice de transition et Distribution stationnaire de chaîne de Markov image

Exécution

```
Donnez l ensemble de depart E:
[1,2,3]

E =

     1     2     3

Donnez l ensemble d arrive F:
[1,2,3]

F =

     1     2     3


e =

     3


f =

     3

f est une fonction injective (bijective),Yn est une chaine de Markov image
Donnez l ensemble des antécédents Xi:
[1,2,3]

X =

     1     2     3

Donnez l ensemble des images Yi associe aux antécédent Xi :
[1, 2,3]

Y =

     1     2     3


Q =

         0    0.5000    0.5000
    0.5000         0         0
    0.5000         0         0


dsQ =

    0.3333    0.3333    0.3333


verif =

    0.3333    0.3333    0.3333

q =

    1.0000


la distribution image est stationnaire
```

FIGURE 3.12 – Exécution

3.3 Critères dérivés pour une chaîne de Markov

3.3.1 Critère de transience

Supposons qu'il existe une fonction V tel que :

$$V(i) = \begin{cases} P(T_1 < \infty) = \sum_{j \in E} P_{i,j} V(j) & \text{si } i \neq a \\ V(a) = 1 & \text{si non} \end{cases}$$

On fixe $a = 1, (a \in E)$

$$\text{Alors } \begin{cases} V(1) = 1 \\ V(2) = \sum_{j \in E} P_{2,j} V(j) \\ V(3) = \sum_{j \in E} P_{3,j} V(j) \end{cases}$$

Aprés les calculs, on trouve :

$$V(i) = [1, 1/3, 2/3]$$

Voici l'algorithme qui permet de vérifier les propriétés de ce critère pour une chaîne de Markov :

L'algorithme

```
function Transience(E,P,V)
%E, V sont des vecteurs lignes
e=length(E);
n1=0;n2=0;n3=0;
for i=1:e
    if V(i)<0
       disp('la chaine n''est pas transience')
       return;
    else
       n1=1;
    end
end
if n1==1
   for i=2:e
       if V(i)==V(1)
          disp('la chaine n''est pas transience')
          return;
       else
           n2=1;
       end
   end
end
```

```
if (n1==1) && (n2==1)
    Esp=[];
    for i=1:e
        for j=1:e
           Esp1=sum(P(i,j)*V(j));
        end
        Esp=[Esp;Esp1];
    end
    for i=1:size(Esp,1)
        if Esp(i)==+inf
            disp('la chaine n''est pas transience')
           return;
        else
           n3=1;
        end
    end
end
```

```
Es=[];
if (n1==1) && (n2==1) && (n3==1)
for i=1:e
      for j=1:e
           Esp3=sum(P(i,j)*V(j))-V(i);
      end
   Es=[Es;Esp3];
 end
    for i=1:size(Es,1)
        if Es(i)>0
             disp('la chaine n''est pas transience')
            return;
        end
    end
end

    disp('la chaine est transience')

end
```

FIGURE 3.13 – L'algorithme qui vérifie les propriétés de critères de transience pour une chaîne de Markov

Exécution

```
>> E=[1 2 3];
>> P=[0,0.5,0.5;0.5,0,0.5;0.5,0.5,0];
>> V=[1 1/3 2/3];
>> Transience(E,P,V)
la chaine est transience
```

FIGURE 3.14 – Exécution

3.4 Critères dérivés pour une chaîne de Markov image

3.4.1 Critère de transience

Supposons qu'il existe une fonction $V : E \to \mathbb{R}$, par conséquent il existe une fonction $W : F \to \mathbb{R}$ tel que : $W(y) = \begin{cases} V(j) & \text{si } A = \{j\} \\ \sum\limits_{j\in A} V(j) \min\limits_{i\in E} \dfrac{P_{i,j}}{\sum\limits_{k\in A} P_{i,k}} & \text{si } CardA > 1 \end{cases}$

Voici l'algorithme qui permet de vérifier les propriétés de ce critère pour une chaîne de Markov image Y_n :

L'algorithme

```
Wx=[];
if e==f %la fonction est injective(bijective)
    W=V;
    disp('la chaine image est transience')
else %la fonction est surjective
    for x=1:e
      for y=1:f
          for k=1:f
           W1=sum(V(y)*min(P(x,y)/sum(P(x,k))));
          end
      end
   Wx=[Wx;W1];
    end
    W=Wx'
m1=0;m2=0;m3=0;
for x=1:e
    if W(x)<0
       disp('la chaine n''est pas transience')
       return;
    else
       m1=1;
    end
end
```

```
if n1==1
   for x=2:e
       if W(x)==W(1)
          disp('la chaine n''est pas transience')
          return;
       else
           m2=1;
       end
   end
end
if (m1==1) && (m2==1)
    Espw=[];
    for x=1:e
        for y=1:f
           Espw1=sum(Q(x,y)*W(y));
        end
        Espw=[Espw;Espw1];
    end
    for i=1:size(Espw,1)
        if Espw(i)==+inf
            disp('la chaine n''est pas transience')
           return;
        else
           m3=1;
        end
```

```
    end
end
Esw=[];
if (m1==1) && (m2==1) && (m3==1)
for x=1:e
      for y=1:f
           Espw3=sum(Q(x,y)*W(y))-W(x);
      end
   Esw=[Esw;Espw3];
 end
    for i=1:size(Esw,1)
        if Esw(i)>0
             disp('la chaine n''est pas transience')
            return;
        end
    end
end

    disp('la chaine est transience')

end

end
```

FIGURE 3.15 – L'algorithme qui vérifie les propriétés de critères de transience pour une chaîne de Markov image

Exécution

```
>> E=[1 2 3];
>> F=[1 2 3];
>> P=[0,0.5,0.5;0.5,0,0.5;0.5,0.5,0];
>> Mu=[1/3,1/3,1/3];
>> X=[1 2 3];
>> Y=[1 2 2];
>> Transienceimage(E,F,P,Mu,V,X,Y)

k =

     2

la chaine possède une DS

R =

    1.0000

f est une fonction injective (bijective),Yn est une chaîne de Markov image
la chaine est transience
la chaine image est transience
```

FIGURE 3.16 – Exécution

Le code Matlab :

```
function Chaine_image(E,F,P,Mu,V,X,Y)
Z=zeros(length(P));
for i=1:length(P)
    for j=1:length(P)
        if P(i,j)>0
            z(i,j)=1;
        end
    end
end
%Irréductible
I=(eye(length(P))+z)^(length(P)-1);
if I>0
    disp('la chaîne de Markov est irréductible');
else
    disp('la chaîne de Markov est réductible ');
end
%périodique
n=10;
if ~isempty(find(diag(P)))
    disp('la chaîne de Markov est apériodique');
else
    I=P;k=0;
    for i=2:n
        I=I*P;
        if ~isempty(find(diag(I)))
            k=k+1;
            MP(k,1:length(P))=zeros(1,length(P));
            MP(k,find(diag(I))>0)=i;
        end

    end
    d=0;
    for i=1:n
        d= min(MP,[],'all');
    end
    d
    disp('la chaîne de Markov est périodique');
end
%ergodique
if P^((length(P)-1)^2+1)>0
    disp('la chaîne de Markov est ergodique');
else
    disp('la chaîne de Markov n'' est pas ergodique ');
end
%E, V sont des vecteurs lignes
e=length(E);
for k=1:100
    Pk=P^k;
```

```
    if Pk>0
        k %la valeur de la puissance laqualle Pk>0
        break
    end
end
Pk
if Pk>0
    disp('la chaine possède une DS ')
    N=1;
else
    disp('la chaine ne possède pas une DS')
end
%calculer la DS de la chaine de Markov X_n
if N==1
    dim=length(Mu); m=9999999;[S,D]=eig(P.');
    S1=inv(S);LP=S*D*S1;DL=D^m;
    DS=S*DL*S1;NL=DS' ;ds=NL(1,1:dim) ;
    R=0;
    for i=1:dim
        R=R+ds(i);
    end
    R %R=1 le programme marche
end
f=length(F);
% determination de la nature de la fonction f
c=0;
if e <= f
    c=1;
    disp('f est une fonction injective (bijective),Yn est une chaîne de Markov image ')
else
    c=2;
    disp('f est une fonction surjective, il faut vérifier la CNS')
end
% vérification du caractère Markovien de Yn dans le cas surjective.
if c==2
    P1=0;
    P2=0;
    H=0;
    for i=1:e
        i=1;
        c=i+1;
        if Y(1)==Y(c)
            P1=P1+P(1,i); %calcul du premier terme.
            P2=P2+P(c,i); %calcul du second terme.
        end
    end
    if P1==P2
        disp('Yn est une chaîne de Markov image.')
        H=1;
    else disp('Yn n est pas une chaîne de Markov image.')
```

```
        return
    end
end
%calculer l image de la chaine
dsQ=0;
if c==1 % si f est injective
    for i=1:e
        for j=1:f
            Q(i,j)=P(Y(i),Y(j));
        end
    end
    for i=1:f
        dsQ(i)=ds(Y(i));
    end
    Q
    dsQ
else if H==1 %si f est surjective
        for i=1:f
            for j=1:f
                if P(i,j)==0
                    Q(i,j)=P(i+1,j);
                else
                    Q(i,j)=P(i,j);
                end
            end
        end
        for i=1:f
            dsQ=dsQ+ds(i);
        end
        dsQ=[ds(i),dsQ];
end
Q
dsQ
end
% calculer la distribution stationnaire image
verif=dsQ*Q;
vf=0;
if verif==dsQ
    vf=1;
end
ver=0;
if verif>=0
    ver=1;
end
q=0;
for i=1:f
    q=q+dsQ(i);
end
if vf==1 && q==1 && ver==1
    disp('la distribution image est stationnaire ');
```

```
else
    disp('la distribution image n'' est pas stationnaire')  ;
end
n1=0;n2=0;n3=0;
for i=1:e
    if V(i)<0
        disp('la chaine n''est pas transience')
        return;
    else
        n1=1;
    end
end
if n1==1
    for i=2:e
        if V(i)==V(1)
            disp('la chaine n''est pas transience')
            return;
        else
            n2=1;
        end
    end
end
if (n1==1) && (n2==1)
    Esp=[];
    for i=1:e
        for j=1:e
            Esp1=sum(P(i,j)*V(j));
        end
        Esp=[Esp;Esp1];
    end
    for i=1:size(Esp,1)
        if Esp(i)==+inf
            disp('la chaine n''est pas transience')
            return;
        else
            n3=1;
        end
    end
end
Es=[];
if (n1==1) && (n2==1) && (n3==1)
    for i=1:e
        for j=1:e
            Esp3=sum(P(i,j)*V(j))-V(i);
        end
        Es=[Es;Esp3];
    end
    for i=1:size(Es,1)
        if Es(i)>0
            disp('la chaine n''est pas transience')
```

```
            return;
        end
    end
end
disp('la chaine est transience')
Wx=[];
if e==f %la fonction est injective(bijective)
    W=V;
    disp('la chaine image est transience')
else %la fonction est surjective
    for x=1:e
        for y=1:f
            for k=1:f
                W1=sum(V(y)*min(P(x,y)/sum(P(x,k))));
            end
        end
        Wx=[Wx;W1];
    end
    W=Wx'
    m1=0;m2=0;m3=0;
    for x=1:e
        if W(x)<0
            disp('la chaine n''est pas transience')
            return;
        else
            m1=1;
        end
    end
    if n1==1
        for x=2:e
            if W(x)==W(1)
                disp('la chaine n''est pas transience')
                return;
            else
                m2=1;
            end
        end
    end
    if (m1==1) && (m2==1)
        Espw=[];
        for x=1:e
            for y=1:f
                Espw1=sum(Q(x,y)*W(y));
            end
            Espw=[Espw;Espw1];
        end
        for i=1:size(Espw,1)
            if Espw(i)==+inf
                disp('la chaine n''est pas transience')
                return;
```

```
            else
                m3=1;
            end
        end
    end
    Esw=[];
    if (m1==1) && (m2==1) && (m3==1)
        for x=1:e
            for y=1:f
                Espw3=sum(Q(x,y)*W(y))-W(x);
            end
            Esw=[Esw;Espw3];
        end
        for i=1:size(Esw,1)
            if Esw(i)>0
                disp('la chaine n''est pas transience')
                return;
            end
        end
    end
    disp('la chaine est transience')

end
end
```

Conclusion générale

Après les études décrites , on peut répondre à quelques questions et fournir des explications sur le sujet.

L'image d'une chaîne de Markov par une fonction f est une chaîne de Markov si elle satisfait les conditions suffisantes qui sont relié au type de la fonction f (injective , surjective , bijective) et a la matrice de transition.

Si f est surjective, donc $(f(X_n))_n$ est une chaîne de Markov cela permet de réduire l'espace d'états .

Si f n'est pas injective et n'est pas surjective la condition de la matrice de transition, alors l'image est une chaîne de Markov.

La chaîne de Markov image possède les mêmes propriétés que la chaîne de Markov (récurrence, transience, comportement asymptotique...).

Après avoir confirmé que les propriétés des critères dérivés pour une chaîne de Markov sont applicables sur chaîne de Markov image avec des conditions dans le cas où la fonction f est surjective, Plus de détail vérifier si la chaîne de Markov satisfait les conditions des critères dérivés, alors la chaîne de Markov image vérifiée ces conditions.

A la fin, La simulation avec le programme Matlab permet de réaliser numériquement les résultats théoriques qu'on a démontrer dans le deuxième chapitre .

Bibliographie

[1] Amaury Lambert, *Théorie de la Mesure et Intégration*. Université Pierre et Marie Curie (Paris 6),Licence de Mathématiques L3,Année 201112.

[2] Arnaud Bodin. *Algèbre*. université de Lille, Cours de mathématique, 2016.

[3] Béatrice de Tilière. *CHAÎNES DE MARKOV*. Université Paris-Est-Créteil ESIPE, INGENIEUR, 1ère année.

[4] Delhoum Zohra Sabrina, *Cours et exercices corrigés en probabilités*,Deuxième année ń Classes préparatoires ż, École supérieure d'économie d'Oran, 2020-2021.

[5] Didier.Piau, *processus markoviene*.univ Joseph Fourier, 2006/2007.

[6] Dominique Bakry, Laure Coutin, Thierry Delmotte. *Martingales et Chaînes de Markov*. Laboratoire de Statistique et Probabilits, Université Paul Sabatier,Toulouse, FRANCE, 10 mai 2004.

[7] Dominique.Foata, et Aimé.Fuchs. *Processus stochastique, processus de poisson, chaînes de Markov et Martingales*. Dunod, Paris, 2004.

[8] E.Pommiés, S.Robin, *Introduction aux chaîne de Markov homogène*. Institut National Agronomique Paris-Grignon, 16 Juin 2004.

[9] Florent.Malrieu, *Préparation à lépreuve de modélisation*. Universite de Rennes 1 , 2012-2013.

[10] Jean.Bérard, *Chaînes de Markov*. Institut Camille Jordan, Lyon 1.

[11] Jean-Christophe Breton, *Probabilités*. Université de Rennes 1,L3 Mathématiques, JanvierAvril 2014.

[12] Jean-François Le Gall, *Intégration, Probabilités et Processus Aléatoires*. Ecole normale supérieure de Paris, Septembre 2006.

[13] Jean-Jacques Ruch, Marie-Line Chabanol. *CHAÎNES DE MARKOV, ESPÉRANCE CONDITIONNELLE*.Université de Bordeaux 1, Préparation à lAgrégation, 2012-2013.

[14] Jean-Yves.DAUXOIS. *Probabilités*. Université de Franche-Comté, Master Recherche de Mathématiques, 2008-2009.

[15] Laurent Garcin. *Ensembles et applications*. MP Dumont dUrville, mars 2022.

[16] Michel.Bonnefont. *Récurrence et Transience*. Université de Bordeaux,Master MIMSE, 2019-2020.

[17] *Modélisation Stochastique*. Université de Marne La Vallée, Cours de Master 2, 2008.

[18] Mustapha Mechab. *Cours dalgèbre*, LMD Sciences et Techniques, Maths1.

[19] Ph. Briand. *Chaînes de Markov à espace détats discrets*. Université Savoie Mont Blanc, LAboratoire de MAthématique, 2020-2021.

[20] TOUCHE Nassim. *Chaînes de Markov à temps discret*. Université Abderrahmane Mira de Béjaia, Cours de Troisième année Mathématiques Appliquées, 2017 - 2018.

[21] Raphaël Deswarte,Raphaël Forien, *Rappels de probabilités*. Pré-rentrée du MVA ENS Cachan, septembre 2016.

[22] Raphael.Lachieze-Rey*, *Chaînes de Markov (et applications)*. M1 MM Paris Descartes, 2020/2021.

[23] S. Lemaire. *Introduction aux chaînes de Markov*. Université Paris-Sud,L3 Math 2012-2013.

[24] Yueyun.Hu,Laurent.Tournier,*Processus Stochastiques à temps discret*.Institut Galilée Université Paris 13,Cours de Macs 2.

[25] Yves.Caumel. *Probabilités et processus stochastiques*. l' Springer-Verlag France, 2011.