

CHAÎNES DE MARKOV

MICHEL BONNEFONT

Master MIMSE Bordeaux

Dans tout ce cours on se placera sur un espace d'états E fini ou dénombrable.

Dans ce chapitre, on va s'intéresser à l'étude de phénomènes aléatoires dépendant du temps. Dans l'étude d'un tel processus aléatoire, ce qui nous intéresse particulièrement c'est de connaître la loi de ce qui va passer dans le futur en tenant compte de toutes les informations obtenues dans le passé. Dans de nombreuses situations, on peut faire l'hypothèse simplificatrice suivante : " *A un instant t donné, le futur du processus ne dépend pas de tout le passé (c'est-à-dire de toute la trajectoire jusqu'à l'instant t) mais juste de la position actuelle à l'instant t .*" Le processus est alors appelé *processus de Markov* et dans le cas où le temps est discret : *chaîne de Markov*.

Exemples :

- (1) Gestion de stock : Dans de nombreuses situations, on peut considérer que le stock à l'instant $n + 1$: S_{n+1} s'écrit $S_{n+1} = S_n + R_{n+1}$ où R_{n+1} est une variable aléatoire qui ne dépend que de la valeur du stock à l'instant n : S_n .
- (2) Prix d'une action : Dans de nombreuses situations, on peut considérer que le prix d'une action à l'instant $n + 1$: A_{n+1} peut s'écrire : $A_{n+1} = A_n(1 + r_{n+1})$ où r_{n+1} est une variable aléatoire qui ne dépend que de la valeur du prix de l'action à l'instant n : A_n .

1. DÉFINITIONS ET PREMIÈRES PROPRIÉTÉS

Définition 1.1. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(X_n)_{n \in \mathbb{N}}$ une suite variables aléatoires définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans E . On dit que $(X_n)_{n \in \mathbb{N}}$ est une **chaîne de Markov** si, pour tout $(n + 1)$ -uplet (x_0, x_1, \dots, x_n) de points de E tel que $\mathbb{P}(\bigcap_{0 \leq j \leq n-1} \{X_j = x_j\}) > 0$ on a

$$(1.1) \quad \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_0 = x_0) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}).$$

La matrice $Q_n = (Q_n(x, y))_{x, y \in E}$ définie par $Q_n(x, y) = \mathbb{P}(X_n = y | X_{n-1} = x)$ est appelée **matrice de transition** de la chaîne de Markov (au temps n).

$Q_n(x, y)$ désigne donc la probabilité de passer de l'état $x \in E$ à l'état $y \in E$ au temps n .

On dit que la chaîne de Markov est **homogène** si la matrice de transition ne dépend pas de n . On note alors Q la matrice de transition. Dans toute la suite, nous ne considérerons que des chaînes de Markov homogènes.

La loi de X_0 est appelée **loi initiale** de la chaîne de Markov.

La propriété (??) s'appelle la propriété de Markov. Elle équivaut à dire que la loi de X_n conditionnellement à $(X_0, X_1, \dots, X_{n-1})$ est en fait la loi de X_n conditionnellement à X_{n-1} . La propriété de Markov signifie bien que le futur (immédiat) ne dépend du passé (toute la trajectoire) qu'à travers le présent (le point actuel).

Exercice 1.2. Un premier exemple : Marche aléatoire sur \mathbb{Z} . Soit $(Z_n)_{n \in \mathbb{N}}$ une suite de v.a. indépendantes et identiquement distribuées (i.i.d.) de loi de Rademacher $\mathcal{R}(p)$, $0 \leq p \leq 1$; c'est-à-dire,

$$\mathbb{P}(Z_n = +1) = p \text{ et } \mathbb{P}(Z_n = -1) = 1 - p.$$

On définit alors $X_n = Z_1 + \dots + Z_n$, $n \geq 0$ et $X_0 = 0$. Montrer que (X_n) est bien une **chaîne de Markov homogène**. Si $p = 1/2$ on parle de **marche aléatoire symétrique**.

Bien sûr, on peut faire la même chose en dimension d . La marche aléatoire symétrique est alors obtenue en prenant

$$\mathbb{P}(Z_i = e_1) = \mathbb{P}(Z_i = -e_1) = \dots = \mathbb{P}(Z_i = e_d) = \mathbb{P}(Z_i = -e_d) = \frac{1}{2d}, i \geq 1,$$

où (e_1, e_2, \dots, e_d) est la base usuelle de \mathbb{Z}^d .

2. MATRICES STOCHASTIQUES

Définition 2.1. Une matrice $P = (P_{i,j})_{i,j \in E}$ est une **matrice stochastique** si elle vérifie :

- (1) $\forall i, j \in E, P_{i,j} \geq 0$
- (2) $\forall i \in E, \sum_{j \in E} P_{i,j} = 1$

Exemple 2.2.

$$(1) \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{pmatrix} \text{ est une matrice stochastique.}$$

- (2) La matrice $P = (P_{i,j})_{i,j \in \mathbb{N}}$ donnée pour tout $i \in \mathbb{N}^*$ par $\begin{cases} P_{i,i+1} = P_{i,i-1} = 1/4 \\ P_{i,i} = 1/2 \end{cases}$ et $P_{0,0} = P_{0,1} = 1/2$ est stochastique.

Proposition 2.3. Si Q est une matrice de transition d'une chaîne de Markov, alors Q est une matrice stochastique.

Preuve. Exercice. ■

Proposition 2.4. Soient P et Q deux matrices stochastiques de même taille alors le produit PQ est aussi une matrice stochastique.

Démonstration. Soient P et Q deux matrices stochastiques alors

- PQ est bien définie, en effet la série à termes positifs $(PQ)_{i,j} = \sum_{l \in E} P_{i,l}Q_{l,j}$ converge : P et Q sont stochastiques donc, pour tout $l \in E$, $Q_{l,j} \leq 1$ et $\sum_{l \in E} P_{i,l} = 1$, donc $\sum_{l \in E} P_{i,l}Q_{l,j} \leq \sum_{l \in E} P_{i,l} = 1$.
- Montrons que PQ est stochastique :

$$\begin{aligned} \sum_{j \in E} (PQ)_{i,j} &= \sum_{j \in E} \sum_{l \in E} P_{i,l}Q_{l,j} = \sum_{l \in E} \sum_{j \in E} P_{i,l}Q_{l,j} \text{ par Fubini-Tonelli,} \\ &= \sum_{l \in E} P_{i,l} \sum_{j \in E} Q_{l,j} = \sum_{l \in E} P_{i,l} = 1. \end{aligned}$$

□

Corollaire 2.5. *Soit Q une matrice stochastique, alors pour tout $n \in \mathbb{N}$, Q^n est aussi une matrice stochastique. (Pour $n = 0$, par convention, on pose $Q^0 := Id$).*

Démonstration. Récurrence immédiate (à faire soi-même). □

Interprétation et représentation : Il est facile d'associer à une matrice stochastique un graphe orienté, pondéré. L'ensemble des sommets du graphe sera l'espace d'états E et les arêtes orientées (pondérées) représenteront les valeurs non nulles de la matrice stochastique.

Dans le cas d'une matrice de transition Q d'une chaîne de Markov homogène, la pondération sur l'arête orientée (x, y) correspond à la probabilité de la chaîne de passer de l'état x à l'état y .

Exercice 2.6. Faire les graphes associés aux deux exemples de l'Exemple ??.

3. LOI AU TEMPS n DE LA CHAÎNE DE MARKOV

Pour le résultat suivant, on considère (exceptionnellement) que la chaîne de Markov n'est pas nécessairement homogène.

Proposition 3.1. *Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans E . Pour tout $n \in \mathbb{N}^*$, pour tous $x, y \in E$, on définit $Q_n(x, y) = \mathbb{P}(X_n = y | X_{n-1} = x)$. Alors on a :*

$$\forall n \in \mathbb{N}, \forall x_0, x_1, \dots, x_n \in E,$$

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_0 = x_0)Q_1(x_0, x_1) \dots Q_n(x_{n-1}, x_n).$$

Démonstration. Montrons ce résultat par récurrence sur n .

- C'est trivial pour $n=0$.
- Supposons le résultat vrai pour $n \geq 0$ alors

$$\begin{aligned} \mathbb{P}(X_{n+1} = x_{n+1}, \dots, X_0 = x_0) &= \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0) \mathbb{P}(X_n = x_n, \dots, X_0 = x_0), \\ &= \mathbb{P}(X_n = x_n, \dots, X_0 = x_0) \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n), \\ &= \mathbb{P}(X_n = x_n, \dots, X_0 = x_0) Q_{n+1}(x_n, x_{n+1}), \\ &= \mathbb{P}(X_0 = x_0) Q_1(x_0, x_1) \dots Q_n(x_{n-1}, x_n) Q_{n+1}(x_n, x_{n+1}) \end{aligned}$$

On a bien démontré le résultat souhaité. \square

Remarque 3.2. La réciproque est vraie : s'il existe une suite de matrices $((Q_n(x, y))_{x, y \in E})_{n \in \mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}, \forall x_0, x_1, \dots, x_n \in E, \\ \mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_0 = x_0)Q_1(x_0, x_1) \dots Q_n(x_{n-1}, x_n)$$

alors $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov.

Exercice 3.3. Le prouver.

Corollaire 3.4. *Sous les mêmes hypothèses : $(X_n)_{n \geq 0}$ une chaîne de Markov non nécessairement homogène et de loi initiale $\mathcal{L}(X_0) := \mu$.*

La loi de X_n au temps n est donnée par la formule : pour $x \in E$

$$(3.1) \quad \mathbb{P}(X_n = x) = \sum_{z_0, z_1, \dots, z_{n-1} \in E} \mu(z_0)Q_1(z_0, z_1) \dots Q_n(z_{n-1}, x).$$

Ici, nous revenons au cas d'une chaîne de Markov *homogène*. Dans ce cas, nous obtenons le corollaire important suivant

Corollaire 3.5. *Soit $(X_n)_{n \geq 0}$ une chaîne de Markov homogène de matrice de transition Q et de loi initiale $\mathcal{L}(X_0) := \mu$. On note Q^n la puissance n -ième de la matrice Q , alors :*

(1) *La loi de X_n au temps n est donnée par la formule :*

$$\begin{aligned} \mathbb{P}(X_n = x) &= \sum_{z_0, z_1, \dots, z_{n-1} \in E} \mu(z_0)Q(z_0, z_1) \dots Q(z_{n-1}, x) \\ &= \sum_{y \in E} \mu(y)Q^n(y, x) \end{aligned}$$

(2) *Pour $x, y \in E$, le terme $Q^n(x, y)$ correspond à la probabilité de passer de l'état x à l'état y en exactement n étapes.*

Remarque 3.6. Il est bien pratique d'écrire les **mesures** comme des **vecteurs lignes** et les **fonctions** comme des **vecteurs colonnes**.

Avec cette convention, la loi de la chaîne de Markov au temps n de loi initiale μ s'écrit simplement :

$$\mathcal{L}(X_n) = \mu P^n$$

c'est-à-dire

$$\mathbb{P}(X_n = x) = (\mu P^n)(x).$$

Avec cette convention, pour calculer l'espérance $\mathbb{E}[f(Z)]$ où Z est de loi ν sur E et f est une fonction (bornée) sur E , on a

$$\mathbb{E}[f(Z)] = \sum_{z \in E} \nu(z)f(z) = \nu f.$$

En particulier, l'espérance $\mathbb{E}[f(X_n)]$ se calcule comme

$$\mathbb{E}[f(X_n)] = \sum_{z \in E} (\mu P^n)(z) f(z) = \mu P^n f.$$

Démonstration. La probabilité de passer de x à y en 2 étapes est égale à

$$\begin{aligned} \mathbb{P}(X_2 = y | X_0 = x) &= \sum_{z \in E} \mathbb{P}(X_1 = z, X_2 = y | X_0 = x) \\ &= \sum_{z \in E} Q(x, z) Q(z, y). \end{aligned}$$

Ce dernier terme n'est rien d'autre que le terme (x, y) de la matrice Q^2 .

Plus généralement, pour $n \geq 2$, la probabilité de passer de x à y en n étapes est égale à

$$\begin{aligned} \mathbb{P}(X_n = y | X_0 = x) &= \sum_{z_1, \dots, z_{n-1} \in E} \mathbb{P}(X_1 = z_1, \dots, X_{n-1} = z_{n-1}, X_n = y | X_0 = x_0) \\ &= \sum_{z_1, \dots, z_{n-1} \in E} Q(x, z_1) Q(z_1, z_2) \dots Q(z_{n-1}, y). \end{aligned}$$

Ce dernier terme n'est rien d'autre que le terme (x, y) de la matrice Q^n .

Pour obtenir la loi de X_n , il nous suffit de déterminer $\mathbb{P}(X_n = x)$ pour tout $x \in E$. On a

$$\begin{aligned} \mathbb{P}(X_n = x) &= \sum_{z_0 \in E} \mathbb{P}(X_0 = z_0, X_n = x) \\ &= \sum_{z_0 \in E} \mathbb{P}(X_0 = z_0) \mathbb{P}(X_n = x | X_0 = z_0) \\ &= \sum_{z_0 \in E} \mu(z_0) Q^n(z_0, x). \end{aligned}$$

□

Remarque 3.7. Dans le cas d'une chaîne de Markov non nécessairement homogène de loi initiale, la loi de X_n s'écrit

$$\mathcal{L}(X_n) = \mu Q_1 Q_2 \dots Q_n.$$

Exercice 3.8. Calculer P^2 pour les deux matrices de l'exemple ?? . Tracer les graphes de P et P^2 dans chacun des cas.

Exercice 3.9. On note a, b et c les états de la chaîne 1 de l'exemple ?? et f la fonction définie par $f(a) = 1, f(b) = 2, f(c) = 3$. On note μ_0 la mesure de probabilité uniforme sur $\{a, b, c\}$. Calculer la loi de la chaîne au temps 1 et au temps 2. Calculer également : $\mathbb{E}[f(X_1)]$ et $\mathbb{E}[f(X_2)]$.

Même question pour la marche aléatoire symétrique sur \mathbb{Z} de loi initiale δ_0 et pour la fonction f définie par $f(x) = x^2, x \in \mathbb{Z}$.

Corollaire 3.10. *En reprenant les hypothèses de la proposition ?? (chaîne non nécessairement homogène), on a également, pour $n \in \mathbb{N}, 0 \leq k \leq n, x_k, x_{k+1}, \dots, x_n \in E$,*

$$\mathbb{P}(X_n = x_n, \dots, X_k = x_k) = \mathbb{P}(X_k = x_k)Q_k(x_k, x_{k+1}) \dots Q_n(x_{n-1}, x_n).$$

Démonstration.

$$\begin{aligned} & \mathbb{P}(X_n = x_n, \dots, X_k = x_k) \\ = & \sum_{x_0, x_1, \dots, x_{k-1} \in E} \mathbb{P}(X_n = x_n, \dots, X_k = x_k, X_{k-1} = x_{k-1}, \dots, X_0 = x_0), \\ = & \sum_{x_0, x_1, \dots, x_{k-1} \in E} \mathbb{P}(X_0 = x_0)Q_1(x_0, x_1) \dots Q_k(x_{k-1}, x_k)Q_{k+1}(x_k, x_{k+1}) \dots Q_n(x_{n-1}, x_n), \\ = & \sum_{x_0, x_1, \dots, x_{k-1} \in E} \mathbb{P}(X_k = x_k, X_{k-1} = x_{k-1}, \dots, X_0 = x_0)Q_{k+1}(x_k, x_{k+1}) \dots Q_n(x_{n-1}, x_n), \\ = & \mathbb{P}(X_k = x_k)Q_{k+1}(x_k, x_{k+1}) \dots Q_n(x_{n-1}, x_n). \end{aligned}$$

□

4. UNE CARACTÉRISATION DES CHAÎNES DE MARKOV

Proposition 4.1. *Soit E et F deux ensembles dénombrables, et pour tout $n \geq 1$, soit f_n une application de $E \times F$ dans E . Soit X_0 et $(Y_n)_{n \geq 1}$ des v.a. mutuellement indépendantes, X_0 à valeurs dans E , Y_n à valeurs dans F pour tout $n \geq 1$. Soit enfin $(X_n)_{n \geq 1}$ à valeurs dans E définies par $X_{n+1} = f_n(X_n, Y_{n+1}), \forall n \in \mathbb{N}$. Alors $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov.*

Démonstration. On remarque que

$$\begin{aligned} \mathbb{P}(X_{n+1} = x_{n+1}, \dots, X_0 = x_0) &= \mathbb{P}(f_n(x_n, Y_{n+1}) = x_{n+1}, X_n = x_n, \dots, X_0 = x_0), \\ &= \sum_{z, f_n(x_n, z) = x_{n+1}} \mathbb{P}(Y_{n+1} = z, X_n = x_n, \dots, X_0 = x_0), \\ &= \sum_{z, f_n(x_n, z) = x_{n+1}} \mathbb{P}(Y_{n+1} = z) \mathbb{P}(X_n = x_n, \dots, X_0 = x_0). \end{aligned}$$

Donc on a

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0) = \sum_{z, f_n(x_n, z) = x_{n+1}} \mathbb{P}(Y_{n+1} = z).$$

De la même façon on peut facilement montrer que

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) = \sum_{z, f_n(x_n, z) = x_{n+1}} \mathbb{P}(Y_{n+1} = z)$$

ce qui montre bien que $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov. □

Remarque 4.2. Si de plus, $f_n = f$ et les $(Y_n)_{n \in \mathbb{N}}$ sont dentiquement distribués et que $X_{n+1} = f(X_n, Y_{n+1})$, alors $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov homogène.

Remarque 4.3. La réciproque est vraie. Si (X_n) est une chaîne de Markov sur E , alors on peut représenter $X_{n+1} = f_n(X_n, Y_{n+1})$ pour une certaine fonction $f : (\mathbb{N}, E, F) \rightarrow E$ et des variables aléatoires indépendantes Y_n et indépendante de X_0 .

5. PROPRIÉTÉS DE MARKOV FAIBLE ET FORTE

5.1. Propriété de Markov faible.

Notation 5.1. Dans toute la suite on notera \mathcal{F}_n la tribu engendrée par les v.a. X_0, \dots, X_n , autrement dit $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$.

On notera également \mathbb{P}_x la probabilité conditionnelle sachant $\{X_0 = x\}$, $x \in E$.

Théorème 5.2. Propriété de Markov faible

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov. Alors pour tous $n \in \mathbb{N}$, $x \in E$ et conditionnellement à $\{X_n = x\}$, $(X_{n+p})_{p \in \mathbb{N}}$ est une chaîne de Markov de loi initiale δ_x . De plus elle est indépendante de \mathcal{F}_n ; c'est-à-dire, pour tout $A \in \mathcal{F}_n$, pour tous $m > 0$, $x_0, x_1, \dots, x_m \in E$

$$\begin{aligned} & \mathbb{P}(A \cap \{X_{n+1} = x_1, \dots, X_{n+m} = x_m\} | X_n = x) \\ &= \mathbb{P}(X_{n+1} = x_1, \dots, X_{n+m} = x_m | X_n = x) \mathbb{P}(A | X_n = x) \\ &= \mathbb{P}_x(X_1 = x_1, \dots, X_m = x_m) \mathbb{P}(A | X_n = x). \end{aligned}$$

Avant de faire la démonstration, nous allons établir 2 lemmes.

Lemme 5.3. *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et soient $A, B, C \in \mathcal{F}$ tels que $\mathbb{P}(B \cap C) > 0$. Alors*

$$\mathbb{P}(A|B|C) = \mathbb{P}(A|B \cap C).$$

Démonstration. Soit $\mathbb{Q}(\cdot) := \mathbb{P}(\cdot|C)$. \mathbb{Q} est une probabilité. Par définition de la probabilité conditionnelle, on a

$$\begin{aligned} \mathbb{P}(A|B|C) &= \mathbb{Q}(A|B) = \frac{\mathbb{Q}(A \cap B)}{\mathbb{Q}(B)} = \frac{\mathbb{P}(A \cap B|C)}{\mathbb{P}(B|C)} \\ &= \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)} = \mathbb{P}(A|B \cap C). \end{aligned}$$

□

Lemme 5.4. *Soit $(X_n)_{n \geq 0}$ une chaîne de Markov. Soit $n \geq 0$ et $x_0, x_n, x_{n+1} \in E$ tels que $\mathbb{P}(\{X_n = x_n\} \cap \{X_0 = x_0\}) > 0$. Alors*

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n).$$

Démonstration. En sommant sur toutes les trajectoires possibles et en utilisant la définition d'une chaîne de Markov, on trouve :

$$\begin{aligned}
& \mathbb{P}(X_{n+1} = x_{n+1}, X_n = x_n, X_0 = x_0) \\
&= \sum_{z_1, z_2, z_{n-1} \in E} \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = z_{n-1}, \dots, X_1 = z_1, X_0 = x_0) \\
&= \sum_{z_1, z_2, z_{n-1} \in E} \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = z_{n-1}, \dots, X_1 = z_1, X_0 = x_0) \\
&\quad \times \mathbb{P}(X_n = x_n, X_{n-1} = z_{n-1}, \dots, X_1 = z_1, X_0 = x_0) \\
&= \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \sum_{z_1, z_2, z_{n-1} \in E} \mathbb{P}(X_n = x_n, X_{n-1} = z_{n-1}, \dots, X_1 = z_1, X_0 = x_0) \\
&= \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \mathbb{P}(X_n = x_n, X_0 = x_0)
\end{aligned}$$

D'où :

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n).$$

□

Revenons à la démonstration de la propriété de Markov faible.

Démonstration. Montrons d'abord que $(X_{n+p})_{p \in \mathbb{N}}$ est une chaîne de Markov. Soit, $m \geq 1$ et $x_1, \dots, x_m \in E$, on a en utilisant le lemme ci-dessus :

$$\begin{aligned}
& \mathbb{P}_x(X_{m+n} = x_m | X_{n+m-1} = x_{m-1}, \dots, X_{n+1} = x_1) \\
&= \mathbb{P}(X_{m+n} = x_m | X_{n+m-1} = x_{m-1}, \dots, X_{n+1} = x_1, X_n = x) \\
&= \mathbb{P}(X_{m+n} = x_m | X_{n+m-1} = x_{m-1}) \\
&= \mathbb{P}(X_{m+n} = x_m | X_{n+m-1} = x_{m-1}, X_n = x) \\
&= \mathbb{P}_x(X_{m+n} = x_m | X_{n+m-1} = x_{m-1}).
\end{aligned}$$

Montrons maintenant l'indépendance conditionnelle. Il suffit de la montrer pour $A = \{X_0 = y_0, \dots, X_n = y_n\}$. En effet par σ -additivité de \mathbb{P} , on peut l'obtenir pour tout $A \in \mathcal{F}_n$. Il suffit de plus de regarder le cas $y_n = x$ (sinon les deux membres sont nuls). On a

$$\begin{aligned}
& \mathbb{P}(A \cap \{X_{n+1} = x_1, \dots, X_{n+m} = x_m\} | X_n = x) \\
&= \mathbb{P}(X_0 = y_0, \dots, X_n = x, X_{n+1} = x_1, \dots, X_{n+m} = x_m | X_n = x), \\
&= \frac{\mathbb{P}(X_0 = y_0, \dots, X_n = x, X_{n+1} = x_1, \dots, X_{n+m} = x_m)}{\mathbb{P}(X_n = x)}, \\
&= \frac{\mathbb{P}(X_0 = y_0, \dots, X_n = x)}{\mathbb{P}(X_n = x)} Q_{n+1}(x, x_1) Q_{n+2}(x_1, x_2) \dots Q_{n+m}(x_{m-1}, x_m), \\
&= \frac{\mathbb{P}(A)}{\mathbb{P}(X_n = x)} \mathbb{P}(X_{n+1} = x_1, \dots, X_{n+m} = x_m | X_n = x) \\
&= \mathbb{P}(A | X_n = x) \mathbb{P}(X_{n+1} = x_1, \dots, X_{n+m} = x_m | X_n = x).
\end{aligned}$$

□

5.2. Propriété de Markov forte.

Définitions 5.5.

- Une **filtration** est une suite croissante de tribus, c'est à dire une suite de tribus $(\mathcal{G}_n)_{n \in \mathbb{N}}$ vérifiant $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \dots \subset \mathcal{G}_n \subset \dots$
- $(\sigma(X_0, \dots, X_n))_{n \in \mathbb{N}} = (\mathcal{F}_n)_{n \in \mathbb{N}}$ est appelée **filtration engendrée** par $(X_n)_{n \in \mathbb{N}}$.

Définition 5.6. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. Soit τ une v.a. définies sur le même espace probabilisé. On dit que τ est un **temps d'arrêt** adapté à la suite $(X_n)_{n \in \mathbb{N}}$ (ou temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$) si :

- a) τ prend ses valeurs dans $\mathbb{N} \cup \{\infty\}$,
- b) $\forall n \geq 0$, on a $\{\tau \leq n\} \in \mathcal{F}_n$.

Remarque 5.7. On peut remplacer la condition b) par la condition

- b') $\forall n \geq 0$, on a $\{\tau = n\} \in \mathcal{F}_n$.

Exercice 5.8. Le prouver !

Exercice 5.9. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. à valeurs dans E . Montrer que les v.a. suivantes sont des temps d'arrêt adaptés à la suite X_n :

- les v.a. constantes p.s.,
- le temps d'entrée dans l'ensemble A , noté T_A , donné par $T_A = \inf\{n \in \mathbb{N}, X_n \in A\}$,
- le temps de retour en A , noté S_A , donné par $S_A = \inf\{n \in \mathbb{N}^*, X_n \in A\}$,
- $T \wedge S$, $T \vee S$ et $T + S$ où T et S sont des temps d'arrêt.

Proposition-Définition 5.10. Soit τ un temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_n$. On pose $\mathcal{F}_\infty = \sigma(\mathcal{F}_n, n \geq 0)$. On définit $\mathcal{F}_\tau \subset \mathcal{F}_\infty$ par $\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty, \forall n \geq 0, A \cap \{\tau = n\} \in \mathcal{F}_n\}$. \mathcal{F}_τ est une tribu appelée **tribu des événements antérieurs** à τ .

Démonstration.

- $\forall n \geq 0, \Omega \cap \{\tau = n\} = \{\tau = n\} \in \mathcal{F}_n$ donc $\Omega \in \mathcal{F}_\tau$.
- Soit $(A_k)_{k \in \mathbb{N}} \in \mathcal{F}_\tau$, montrons que $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{F}_\tau$. On a

$$\left(\bigcup_{k \in \mathbb{N}} A_k \right) \cap \{\tau \leq n\} = \bigcup_{k \in \mathbb{N}} (A_k \cap \{\tau \leq n\}) \in \mathcal{F}_n.$$

Donc $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{F}_\tau$.

- Soit $A \in \mathcal{F}_\tau$, montrons que $\bar{A} \in \mathcal{F}_\tau$. Soit $B = \bar{A} \cap \{\tau \leq n\}$, on a

$$\begin{aligned} \bar{A} \cap \{\tau \leq n\} &= (\bar{A} \cup \{\tau \leq n\}) \cap \{\tau \leq n\} \\ &= \overline{(A \cap \{\tau \leq n\})} \cap \{\tau \leq n\} \in \mathcal{F}_n, \end{aligned}$$

car $A \cap \{\tau \leq n\}$ et $\{\tau \leq n\}$ sont dans \mathcal{F}_n . Donc $\bar{A} \in \mathcal{F}_\tau$. □

Exercice 5.11. On reprend l'exemple de la marche aléatoire sur \mathbb{Z} . On pose $\tau := \inf\{n, X_n = -10\}$. On note $B := \{\exists n \in \mathbb{N}, n \leq \tau, X_n = 10\}$. Montrer que B appartient à \mathcal{F}_τ .

Exercice 5.12. Soit τ un temps d'arrêt. On note $\sigma(\tau)$ la tribu engendrée par τ . Montrer que l'on a l'inclusion : $\sigma(T) \subset \mathcal{F}_\tau$. A-t-on égalité ?

Théorème 5.13. Propriété de Markov forte

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov **homogène** de paramètres (μ, P) et τ un temps d'arrêt adapté à $(X_n)_{n \in \mathbb{N}}$. Pour tout $x \in E$ et conditionnellement à $\{\tau < \infty, X_\tau = x\}$, la suite $(X_{\tau+n})_{n \in \mathbb{N}}$ est une chaîne de Markov de paramètres (δ_x, P) . De plus elle est indépendante de \mathcal{F}_τ ; c'est-à-dire, pour tout $A \in \mathcal{F}_\tau$, $x_1, \dots, x_m \in E$

$$\begin{aligned} & \mathbb{P}(\{X_{\tau+n} = x_n, \dots, X_\tau = x_0\} \cap A | \{\tau < +\infty\} \cap X_\tau = x_0) \\ &= \mathbb{P}_{x_0}(\{X_n = x_n, \dots, X_1 = x\}) \mathbb{P}(A | \{\tau < +\infty\} \cap X_\tau = x_0). \end{aligned}$$

Remarque 5.14. En prenant $\tau = k$ on retrouve la propriété de Markov faible.

Remarque 5.15. En particulier, on a aussi que

$$\begin{aligned} & \mathbb{P}(\{X_{\tau+n} = x_n, \dots, X_\tau = x_0\} | \{\tau < +\infty\} \cap X_\tau = x_0) \\ &= \mathbb{P}_{x_0}(\{X_n = x_n, \dots, X_1 = x\}). \end{aligned}$$

Démonstration. Soit $A \in \mathcal{F}_\tau$, en écrivant $A \cap \{\tau < \infty\}$ comme l'union disjointe $A \cap \{\tau < \infty\} = \bigcup_{m \geq 0} A \cap \{\tau = m\}$ on a

$$\begin{aligned} & \mathbb{P}(\{X_{\tau+n} = x_n, \dots, X_\tau = x_0\} \cap A \cap \{\tau < +\infty\}) \\ &= \sum_{m \geq 0} \mathbb{P}(\{X_{\tau+n} = x_n, \dots, X_\tau = x_0\} \cap A \cap \{\tau = m\} \cap \{X_\tau = x_0\}) \\ &= \sum_{m \geq 0} \mathbb{P}(\{X_{m+n} = x_n, \dots, X_m = x_0\} \cap A \cap \{\tau = m\} \cap \{X_m = x\}) \\ &= \sum_{m \geq 0} \mathbb{P}(\{X_{m+n} = x_n, \dots, X_m = x_0\} | X_m = x) \mathbb{P}(A \cap \{\tau = m\} | \{X_m = x\}) \mathbb{P}(\{X_m = x\}) \\ &= \mathbb{P}_{x_0}(\{X_n = x_n, \dots, X_1 = x\}) \sum_{m \geq 0} \mathbb{P}(A \cap \{\tau = m\} \cap \{X_m = x_0\}) \\ &= \mathbb{P}_{x_0}(\{X_n = x_n, \dots, X_1 = x\}) \mathbb{P}(A \cap \{\tau < +\infty\} \cap \{X_\tau = x_0\}); \end{aligned}$$

où on a utilisé la propriété de Markov faible à la troisième égalité en remarquant que $A \cap \{\tau = m\} \in \mathcal{F}_m$. Donc, pour $A \in \mathcal{F}_\tau$, en divisant par $\mathbb{P}(\{\tau < +\infty\} \cap X_\tau = x_0)$, on obtient

$$\begin{aligned} & \mathbb{P}(\{X_{\tau+n} = x_n, \dots, X_\tau = x_0\} \cap A | \{\tau < +\infty\} \cap \{X_\tau = x_0\}) \\ &= \mathbb{P}_{x_0}(\{X_n = x_n, \dots, X_1 = x\}) \mathbb{P}(A | \{\tau < +\infty\} \cap X_\tau = x_0). \end{aligned}$$

En particulier, on remarque que

$$\begin{aligned} & \mathbb{P}(\{X_{\tau+n} = x_n, \dots, X_\tau = x_0\} | \{\tau < +\infty\} \cap X_\tau = x_0) \\ &= \mathbb{P}_{x_0}(\{X_n = x_n, \dots, X_1 = x\}). \end{aligned}$$

et donc en particulier, on remarque que

$$\begin{aligned}
 & \mathbb{P}_{\{\tau < +\infty \cap X_\tau = x_0\}}(X_{\tau+n+1} = x_{n+1} | X_{\tau+n} = x_n \dots, X_\tau = x_0) \\
 &= \mathbb{P}_{x_0}(X_{n+1} = x_{n+1} | X_n = x_n \dots, X_0 = x_0 \dots, X_1 = x) \\
 &= \mathbb{P}_{x_0}(X_{n+1} = x_{n+1} | X_n = x_n) \\
 &= \mathbb{P}_{\{\tau < +\infty \cap X_\tau = x_0\}}(X_{\tau+n+1} = x_{n+1} | X_{\tau+n} = x_n).
 \end{aligned}$$

Cette dernière égalité traduit exactement la propriété de Markov de la suite $(X_{\tau+n})_{n \in \mathbb{N}}$ conditionnellement à l'évènement $\{\tau < \infty, X_\tau = x\}$. \square

Exemple d'application : Soit $(X_n)_{n \geq 0}$ une marche aléatoire sur \mathbb{Z} partant de 0. Soit τ le temps d'atteinte de 10 :

$$\tau := \inf\{k \geq 0, X_k = 10\}.$$

Alors conditionnellement à $\{\tau < +\infty\}$, $(X_{\tau+n})_{n \geq 0}$ est aussi une marche aléatoire partant de 10. De plus elle est indépendante de \mathcal{F}_τ , c'est-à-dire elle est indépendante des trajectoires avant τ .