



18

PROBABILITÉS

CHAÎNES DE MARKOV

POUR BIEN DÉMARRER...

1. Revoir le chapitre sur les graphes !
2. On classifie les étudiants de CPGE ECG en trois groupes :
 - Groupe 1 : étudiants qui ne travaillent pas ;
 - Groupe 2 : étudiants qui travaillent mais avec des résultats encore faibles ;
 - Groupe 3 : étudiants qui travaillent et avec des résultats très encourageants.

On choisit un étudiant au hasard en CPGE ECG et on note, pour tout $n \in \mathbb{N}^*$, X_n la variable aléatoire égale au groupe dans lequel se situe l'étudiant après n semaines passées en CPGE. On considère $X_0 = 1$.

On considère que, d'une semaine à l'autre :

- 20% des étudiants du groupe 1 passent dans le groupe 2, les autres restent dans le groupe 1 ;
- 20% des étudiants du groupe 2 passent dans le groupe 3, 10% retournent dans le groupe 1 et les autres restent dans le groupe 2 ;
- 10% des étudiants du groupe 3 retournent dans le groupe 2 et les autres restent dans le groupe 3.

On note, pour tout $n \in \mathbb{N}$, $U_n = (\mathbb{P}(X_n = 1) \quad \mathbb{P}(X_n = 2) \quad \mathbb{P}(X_n = 3))$.

- Représenter la situation par un graphe pondéré et orienté.

- On admet que pour tout $n \in \llbracket 2; +\infty \rrbracket$, les probabilités $\mathbb{P}(X_n = 1)$, $\mathbb{P}(X_n = 2)$ et $\mathbb{P}(X_n = 3)$ sont non nulles. Exprimer, pour tout $n \in \llbracket 2; +\infty \rrbracket$, la probabilité $\mathbb{P}(X_{n+1} = 1)$ en fonction de $\mathbb{P}(X_n = 1)$, $\mathbb{P}(X_n = 2)$ et $\mathbb{P}(X_n = 3)$.

- En déduire une matrice M telle que : $\forall n \in \mathbb{N}, U_{n+1} = U_n M$. Qu'en déduire ?

- Que peut-on dire de la matrice M ?

I GRAPHE PROBABILISTE

DÉFINITION 1

GRAPHE ORIENTÉ ET PONDÉRÉ

Un **graphe orienté et pondéré** $\mathcal{G}$ est la donnée d'un triplet $(S, \mathcal{A}, p)$, où :

- S est un ensemble fini, appelé ensemble des **sommets** de $\mathcal{G}$,
- $\mathcal{A}$ est un ensemble de *couples* de sommets, appelé ensemble des **arcs** de $\mathcal{G}$,
- p est une application de $\mathcal{A}$ dans $\mathbb{R}$, qui à chaque arc associe un réel appelé **poids**.

EXEMPLES 1

E1 Le graphe de l'exemple introductif est un graphe orienté et pondéré.

E2 Une carte routière peut se modéliser par un graphe orienté et pondéré dont les sommets sont les villes et les arêtes ou arcs sont les axes routiers indiquant les distances à parcourir sur chaque axe.

DÉFINITION 2

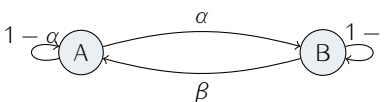
GRAPHE PROBABILISTE

Un **graphe probabiliste** est un graphe orienté et pondéré $(S, \mathcal{A}, p)$ tel que :

- $\forall (i, j) \in \mathcal{A}, p(i, j) \geq 0$ (les poids sont positifs),
- $\forall i \in S, \sum_{j \in S, (i, j) \in \mathcal{A}} p(i, j) = 1$ (la somme des poids issus d'un même sommet est égale à 1).

EXEMPLES 2

E1 Le graphe de l'exemple introductif est un graphe probabiliste.

E2 Pour tous $\alpha, \beta \in [0, 1]$, le graphe  est un graphe probabiliste.

Rappel...

La **matrice d'adjacence** d'un graphe $\mathcal{G}$ d'ordre n est la matrice $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ telle que pour tous $i, j \in \llbracket 1; n \rrbracket$:

- $m_{i,j}$ est le nombre d'arêtes reliant s_i et s_j si $\mathcal{G}$ est non orienté;
- $m_{i,j}$ est le nombre d'arcs de s_i vers s_j si $\mathcal{G}$ est orienté.

Dans le cas des graphes pondérés et probabilistes, on ne parle pas de matrice d'adjacence... Mais il y a bien une matrice associée à de tels graphes. Nous en parlons dans la prochaine partie.

II CHAÎNE DE MARKOV

DÉFINITION 3

CHAÎNE DE MARKOV

Soient E un ensemble fini et $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires toutes définies sur le même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

La suite $(X_n)_{n \in \mathbb{N}}$ est une **chaîne de Markov** sur E lorsque :

- ✓ $\forall n \in \mathbb{N}, X_n(\Omega) \subset E$,
- ✓ pour tout $n \in \mathbb{N}$, pour tout $(e_0, \dots, e_{n+1}) \in E^{n+2}$ tel que $\mathbb{P}([X_0 = e_0] \cap \dots \cap [X_n = e_n]) \neq 0$, on a :

$$\mathbb{P}_{[X_0=e_0] \cap \dots \cap [X_n=e_n]}([X_{n+1} = e_{n+1}]) = \mathbb{P}_{[X_n=e_n]}([X_{n+1} = e_{n+1}])$$

Bien souvent, n désigne un instant et X_n une position ou un état à l'instant n .

Pourquoi ?

Pourquoi, sous l'hypothèse, a-t-on également $\mathbb{P}([X_n = e_n]) \neq 0$?

En gros...

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov lorsque pour tout $n \in \mathbb{N}$, la loi de X_{n+1} ne dépend que de la loi de X_n , et pas de la loi des $X_0, \dots, X_{n-1}$.

Et dans ce cas, $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov lorsque pour tout $n \in \mathbb{N}$, l'état à l'instant $n+1$ ne dépend que de l'état à l'instant n (et pas aux instants précédents).

DÉFINITION 4

CHAÎNE DE MARKOV HOMOGÈNE

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov sur un ensemble fini E .

La chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ est **homogène** lorsque pour tout $n \in \mathbb{N}$, pour tout $(a, b) \in E^2$, la probabilité $\mathbb{P}_{[X_n=a]}([X_{n+1} = b])$ ne dépend pas de n (et donc : $\mathbb{P}_{[X_n=a]}([X_{n+1} = b]) = \mathbb{P}_{[X_0=a]}([X_1 = b])$).

EXEMPLE 3

La suite $(X_n)_{n \in \mathbb{N}}$ de l'exemple introductif est une chaîne de Markov homogène sur $\{1; 2; 3\}$.

Dans toute la suite du cours, on notera r un entier naturel supérieur ou égal à 2 et on ne considérera que des chaînes de Markov finies homogènes sur l'ensemble $\llbracket 1; r \rrbracket$.

Pour info...

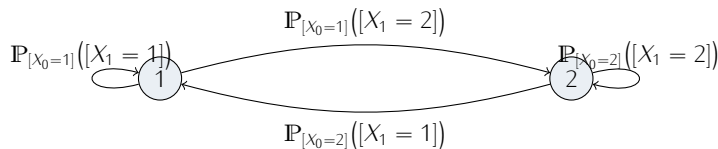
Lorsque l'on considère une *suite* de VA indexée sur un ensemble non dénombrable, par exemple $(X_t)_{t \in \mathbb{R}^+}$, $(X_t)_{t \in [0,1]}$ ou $(X_t)_{t \in \mathbb{R}}$, on parle de **processus aléatoire**. Et on peut, de la même façon, définir un processus aléatoire de Markov fini et homogène (voir ESSEC 2023 Appli 2).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène sur $\llbracket 1; r \rrbracket$.

Le **graphe probabiliste associé à la chaîne de Markov** $(X_n)_{n \in \mathbb{N}}$ est le graphe $\mathcal{G} = (S, \mathcal{A}, p)$ où :

- $S = \llbracket 1; r \rrbracket$,
- $\mathcal{A} = \llbracket 1; r \rrbracket \times \llbracket 1; r \rrbracket$,
- pour tout $(i, j) \in \mathcal{A}$, $p(i, j) = \mathbb{P}_{[X_0=i]}([X_1 = j])$.

Cas particulier $r = 2$:



Important !

Il s'agit bien d'un graphe probabiliste car :

- les poids sont positifs (des probabilités),
- puisque pour tout i , l'application $\mathbb{P}_{[X_0=i]}$ est une probabilité et que $([X_1 = j])_{j \in \llbracket 1; r \rrbracket}$ est un SCE, on a bien $\sum_{j=1}^r \mathbb{P}_{[X_0=i]}([X_1 = j]) = 1$.

Remarque

Et on imagine assez bien ce que cela donne dans le cas $r = 3$ par exemple.

II.1 MATRICE DE TRANSITION

DÉFINITION 6

MATRICE STOCHASTIQUE (PAR LIGNE)

Soit $M = (m_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$. On dit que M est **stochastique (par ligne)** lorsque :

- ✓ $\forall (i, j) \in \llbracket 1; n \rrbracket^2$, $m_{i,j} \geq 0$ (tous les coefficients de M sont positifs),
- ✓ $\forall i \in \llbracket 1; n \rrbracket$, $\sum_{j=1}^n m_{i,j} = 1$ (la somme des coefficients de chaque ligne de M est égale à 1).

Rappel...

Si besoin, on retrace la QUESTION CLASSIQUE 15.

DÉFINITION 7

MATRICE DE TRANSITION

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène sur $\llbracket 1; r \rrbracket$.

La **matrice de transition de** $(X_n)_{n \in \mathbb{N}}$ est la matrice $M = (m_{i,j})_{1 \leq i,j \leq r} \in \mathcal{M}_r(\mathbb{R})$ telle que :

$$\forall (i, j) \in \llbracket 1; r \rrbracket^2, m_{i,j} = \mathbb{P}_{[X_0=i]}([X_1 = j])$$

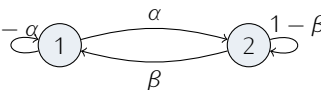
Autrement dit, le coefficient (i, j) de M est la **probabilité de transition** de i vers j .

En gros...

La matrice de transition est une matrice d'adjacence du graphe probabiliste associé à la chaîne de Markov dans laquelle chaque coefficient n'est pas le nombre d'arcs allant de i vers j , mais le poids de l'arc.

EXEMPLES 4

E1 La matrice de transition de la chaîne de Markov dont le graphe probabiliste est



est la matrice

E2 La matrice de transition de l'exemple introductif est la matrice $\begin{pmatrix} 0,8 & 0,2 & 0 \\ 0,1 & 0,7 & 0,2 \\ 0 & 0,1 & 0,9 \end{pmatrix}$.

PROPRIÉTÉ 1

La matrice de transition d'une chaîne de Markov est une matrice stochastique.

Pour info...

Réciproquement, une matrice stochastique peut toujours être associée à un graphe probabiliste et donc à une chaîne de Markov : on peut donc définir un graphe probabiliste par la donnée de sa matrice de transition.

★ DÉMONSTRATION :

Rappel...

Si $\mathbb{P}(A) \neq 0$, alors l'application $\mathbb{P}_A$ est une probabilité. Et donc, si $(B_k)_{k \in \mathbb{N}}$ est un SCE, alors : $\sum_{k=0}^{+\infty} \mathbb{P}_A(B_k) = 1$.

DÉFINITION 8

ÉTAT D'UNE CHAÎNE DE MARKOV

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène sur $\llbracket 1; r \rrbracket$.
Pour tout $n \in \mathbb{N}$, le **n -ième état probabiliste de la chaîne de Markov** $(X_n)_{n \in \mathbb{N}}$ est la matrice ligne, notée U_n , définie par :

$$U_n = (\mathbb{P}([X_n = 1] \quad \dots \quad \mathbb{P}([X_n = r]))$$

Autrement dit :
C'est la matrice ligne composée de la loi de la variable aléatoire X_n .

PROPRIÉTÉ 2

Pour tout $n \in \mathbb{N}$, le n -ème état de $(X_n)_{n \in \mathbb{N}}$ est une matrice ligne stochastique.

★ DÉMONSTRATION : Immédiat car, pour tout $n \in \mathbb{N}$, puisque $X_n(\Omega) \subset \llbracket 1; r \rrbracket$, on a :

$$\forall j \in \llbracket 1; r \rrbracket, \mathbb{P}([X_n = j]) \geq 0 \quad ; \quad \sum_{j=1}^r \mathbb{P}([X_n = j]) = 1$$

★

PROPRIÉTÉ 3

ÉVOLUTION EN TEMPS

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène sur $\llbracket 1; r \rrbracket$.
On note M la matrice de transition associée et, pour tout $n \in \mathbb{N}$, U_n le n -ième état probabiliste de la chaîne.
On a :

$$\forall n \in \mathbb{N}, \quad U_{n+1} = U_n \times M$$

Et donc :

$$\forall n \in \mathbb{N}, \quad U_n = U_0 \times M^n$$

⚠ Attention !
Habituellement, on travaillait sur des matrices colonnes... Ici, on travaille sur des matrices lignes. Dans tous les cas, on vérifie bien la concordance des lignes/colonnes afin que le produit matriciel soit licite.

★ DÉMONSTRATION :

Important !
L'énoncé demande parfois de redémontrer ces deux résultats dans un cas particulier.

★

DÉFINITION 9

ÉTAT STABLE

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène sur $\llbracket 1; r \rrbracket$ dont la matrice de transition est notée M .
Un **état stable** (ou **état invariant** ou **état stationnaire**) de $(X_n)_{n \in \mathbb{N}}$ est une matrice ligne U telle que :

- ✓ U est stochastique,
- ✓ $U \times M = U$

Vocabulaire

On parle parfois de loi de probabilité invariante par M .

EXEMPLE 5

Soit $(X_n)_{n \in \mathbb{N}}$ un chaîne de Markov de matrice de transition M et dont le n -ième état stable est noté U_n . Que dire si U_0 est un état stable ?

♣ **MÉTHODE 1** ♣ Pour déterminer les états stables : on résout, pour $X = (x_1 \dots x_r)$ avec $x_1, \dots, x_r \geq 0$, le système

$$\left\{ \begin{array}{l} XM = X \\ \sum_{i=1}^r x_i = 1 \end{array} \right. \text{ qui est équivalent à } \left\{ \begin{array}{l} {}^tM \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \\ \sum_{i=1}^r x_i = 1 \end{array} \right.$$

Autrement dit :

Les états stables sont les stochastiques U telles que tU est vecteur propre de tM pour la valeur propre 1.

Important !

Puisque M est stochastique, 1 est valeur propre de M et donc 1 est également valeur propre de tM .

THÉORÈME 1

CONVERGENCE EN LOI D'UNE CHAÎNE DE MARKOV (HP ?)

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène sur $\llbracket 1; r \rrbracket$ dont la matrice de transition est notée M .
Si $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une variable aléatoire X , alors la matrice ligne $(\mathbb{P}([X = 1]) \dots \mathbb{P}([X = r]))$ est un état stable de $(X_n)_{n \in \mathbb{N}}$.

Autrement dit :

Si une chaîne de Markov est convergente (en loi) alors elle converge vers un état stable.

Pour info...

Il existe des conditions suffisantes, portant sur M , garantissant l'existence, voire l'unicité d'un état stable ; et même la convergence de $(X_n)_{n \in \mathbb{N}}$ vers cet état stable ! Voir, par exemple, la partie 2 du sujet HEC-ESSEC 2025 Appli.

★ DÉMONSTRATION :

EXEMPLE 6

Reprenons la chaîne de Markov de l'introduction dont la matrice de transition est la matrice $M = \begin{pmatrix} 0,8 & 0,2 & 0 \\ 0,1 & 0,7 & 0,2 \\ 0 & 0,1 & 0,9 \end{pmatrix}$.

- Montrons que $(X_n)_{n \in \mathbb{N}}$ admet un unique état stable et déterminons-le.
- Écrivons une fonction **Python** permettant de calculer et représenter les états 0 à N de $(X_n)_{n \in \mathbb{N}}$.

Pour information, on obtient :

