13

Probabilités

Chaînes de Markov

Introduction...

Pour bien démarrer...

1 # Revoir le vocabulaire sur les graphes.

2 # Méthodes pour calculer des puissances de matrices ?

3 # 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 $n \in \mathbb{N}$, $U_n = \begin{pmatrix} \mathbb{P}([X_n = 1]) & \mathbb{P}([X_n = 2]) & \mathbb{P}([X_n = 3]) \end{pmatrix}$.

1. Représenter la situation par un graphe pondéré et orienté.

2. Soit $k \in [\![1;3]\!]$. Exprimer, pour tout $n \in \mathbb{N}^*$, la probabilité $\mathbb{P}([X_{n+1} = k])$ en fonction de $\mathbb{P}([X_n = 1])$, $\mathbb{P}([X_n = 2])$ et $\mathbb{P}([X_n = 3])$.

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

4. Que peut-on dire de la matrice M ?

I Graphe probabiliste

Définition 1 – Graphe orienté pondéré

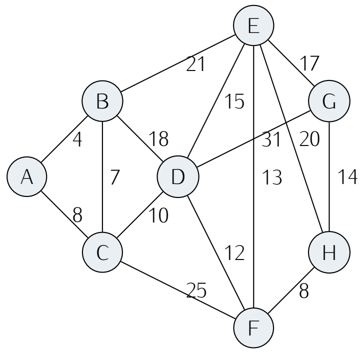
Soit S un ensemble fini.
Un **graphe orienté pondéré** $\mathcal{G}$ est la donnée d'un triplet $(S, \mathcal{A}, v)$, où :

- 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 $\mathcal{G}$,
- v est une application de $\mathcal{A}$ dans $\mathbb{R}$, qui à chaque arc associe un réel appelé **poids**.

Exemples 1

E1 Une carte routière peut se modéliser par un graphe pondéré dont les sommets sont les villes et les arêtes ou arcs sont les axes routiers indiquant les distances à parcourir sur chaque axe.

E2 Le graphe suivant est pondéré, par des poids positifs.



Définition 2 – Graphe probabiliste

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

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

☞ Rappel...

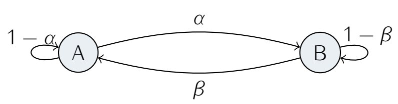
La **matrice d'adjacence** d'un graphe $\mathcal{G}$ d'ordre n est la matrice $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ telle que pour tous $i, j \in [\![1; n]\!]$:
- $m_{i,j}$ est le nombre d'arêtes reliant s_i et s_j si $\mathcal{G}$ est non orienté ;
- $m_{i,j}$ est le nombre d'arcs de s_i vers 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.

Exemples 2

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

E2 Pour tout $\alpha, \beta \in [0; 1]$, le graphe suivant est un graphe probabiliste :



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 $(\Omega, \mathcal{A}, \mathbb{P})$.
La suite $(X_n)_{n \in \mathbb{N}}$ est une **chaîne de Markov sur E** lorsque, pour tout $n \in \mathbb{N}$, pour tout $(e_0, \ldots, e_{n+1}) \in E^{n+2}$:

$$\mathbb{P}\big([X_0 = e_0] \cap \ldots \cap [X_n = e_n]\big) \neq 0 \Longrightarrow \mathbb{P}_{[X_0 = e_0] \cap \ldots \cap [X_n = e_n]}([X_{n+1} = e_{n+1}]) = \mathbb{P}_{[X_n = e_n]}([X_{n+1} = e_{n+1}])$$

Pourquoi ?

Pourquoi $\mathbb{P}([X_n = e_n]) \neq 0$?

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 $X_0, \ldots, X_{n-1}$.

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

En gros...

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 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 $\{1;2;3\}$.

> ☞ **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 E2).

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 $[\![1;r]\!]$.

A toute chaîne de Markov homogène sur $[\![1;r]\!]$, on peut associer un unique graphe probabiliste...
Cas particulier $r = 2$:

II.1 Matrice de transition

Définition 5 - Matrice stochastique (par ligne)

Une matrice est **stochastique (par ligne)** lorsque :

- tous ses coefficients sont positifs,
- la somme des coefficients de chaque ligne est égale à 1.

Propriété 1 (HP - à redémontrer)

Soit $A \in \mathcal{M}_n(\mathbb{R})$.
Si A est stochastique, alors 1 est valeur propre de A et donc de ${}^tA$.

⋆ Démonstration :

⋆

Exemple 4

Écrivons une fonction **Python** prenant en argument d'entrée une matrice A et renvoyant **True** si la matrice est stochastique, et **False** sinon.

> ☞ **Rappel...**
> Si A est un tableau **Python**, la commande **np.shape(A)** renvoie le couple (n, p), ou n est le nombre de lignes et p le nombre de colonnes de A.

Définition 6 - Matrice de transition

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$.
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 :

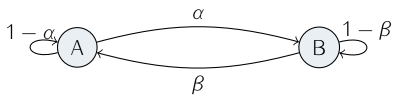
$$\forall (i,j) \in [\![1;r]\!]^2, \;\; m_{i,j} = \mathbb{P}_{[X_0=i]}([X_1=j])$$

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 5

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



est la matrice :

E2 Cas particulier $r = 3$:

Propriété 2

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.

⋆ 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 : $\displaystyle\sum_{k=0}^{+\infty} \mathbb{P}_A(B_k) = 1$.

II.2 Loi d'une chaîne de Markov

Définition 7 - État d'une chaîne de Markov

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$.
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 :

$$U_n = \begin{pmatrix} \mathbb{P}([X_n = 1] & \dots & \mathbb{P}([X_n = r]) \end{pmatrix}$$

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

Propriété 3

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 $X_n(\Omega) \subset [\![1;r]\!]$, on a :

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

Propriété 4 - Évolution en temps

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$.
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 :

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

Et donc :

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

✘ 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.

II.3 Comportement limite

Vocabulaire
On parle parfois de loi de probabilité invariante par M.

Définition 8 – État stable

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$ 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,
- $U = U \times M$

Exemple 6

Que dire de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ si X_0 est un état stable de la matrice de transition ?

Théorème 1 – Existence d'états stables

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$ dont la matrice de transition est notée M.

T1# Si U est un état stable de $(X_n)_{n\in\mathbb{N}}$, alors ${}^tU$ est vecteur propre de ${}^tM$ associé à la valeur propre 1.

T2# Il existe au moins une matrice ligne U telle que U est stochastique et $U = U \times M$.

Petite remarque
T2 est hors programme.

Démonstration :

T1# Supposons que U est un état stable de $(X_n)_{n\in\mathbb{N}}$. Ainsi :

$$U = U \times M$$

D'où :

$${}^tU = {}^t(UM) = {}^tM{}^tU$$

D'où le résultat.

T2# Démonstration loin d'être triviale...

Théorème 2 – Convergence en loi d'une chaîne de Markov

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$ 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 $\begin{pmatrix}\mathbb{P}([X=1]) & \dots & \mathbb{P}([X=r])\end{pmatrix}$ 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.

Démonstration :

Exemple 7

Reprenons la chaîne de Markov de l'introduction.
Montrons qu'elle admet un unique état stable et déterminons-le. Étudions la convergence en loi de $(X_n)_{n\in\mathbb{N}}$.

Finissons sur un théorème hors programme qui fait le lien entre les différentes notions déjà vues.

Théorème 3 (HP)

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur $[\![1;r]\!]$ dont la matrice de transition est notée M.
S'il existe $k \in \mathbb{N}^*$ tel que tous les coefficients de M^k sont strictement positifs, alors :

- la matrice M possède un unique état stable,
- $(X_n)_{n\in\mathbb{N}}$ converge en loi vers l'unique état stable.

☞ Pour info...
C'est un cas particulier du théorème de Perron-Frobénius.

C'est parfait !
C'est le cas si le graphe orienté est fortement connexe ; autrement si, s'il existe toujours un circuit permettant de passer d'un sommet à un autre.
C'est en particulier le cas si la matrice de transition est à coefficients strictement positifs.

⋆ Démonstration : En exercice. ⋆