Utilisation des matrices et chaînes de Markov

Compétences attendues en fin de chapitre :

- Associer un graphe pondéré à une chaîne de Markov.
- Étude de graphe Eulérien.
- Étudier une chaîne de Markov à deux ou trois états.

Le mathématicien du chapitre :

Andreï Markov (1856-1922) est un mathématicien russe considéré comme le père des processus stochastiques. C'est en 1902 qu'il introduit la notion de chaîne étudiée ici. Elles modélisent des phénomènes dynamiques aléatoires dans lesquels le passé n'intervient que via le dernier instant de la chaîne. En 1906, on lui doit la démonstration du théorème limite centrale, très fréquemment utilisé en probabilité de nos jours.

1) Suites de matrices

a) Suites de matrices colonnes

On va mélanger ici deux notions sans que la difficulté soit augmentée.

Définition :

Soit n un entier naturel. On appelle suite de matrices colonnes, notée (U_n), des matrices colonnes dont tous les éléments sont des termes de suites numériques.

Exemple

On donne, $\forall n \in \mathbb{N}, U_n = \begin{pmatrix} 5+n \\ -n^2 \\ 2n-1 \end{pmatrix}$. (U_n) est une suite de matrices colonnes à trois lignes dont les coefficients sont les suites numériques (u_n), (v_n) et (w_n)

On pourrait définir de la même manière une suite de matrices lignes

Définition :

On dit qu'une suite de matrices notée (U_n) converge si et seulement si tous les éléments convergent. La limite de (U_n) est alors la matrice ayant pour coefficients les limites de chaque terme de (U_n).

Exemple :

- On donne $U_n = \begin{pmatrix} n^2 \\ 1-\frac{3}{n} \end{pmatrix}$. (U_n) est divergente car le premier terme tend vers $+\infty$
- On donne $V_n = \begin{pmatrix} -5 \\ 4+\frac{3}{n^2} \end{pmatrix}$. (V_n) est convergente. Sa limite est $L = \begin{pmatrix} -5 \\ 4 \end{pmatrix}$

b) Suite de matrices définies par une relation de récurrence

Exemple introductif :

En 2019, on compte dans un pays 65 millions de citadins et 15 millions de ruraux. On suppose que la population de ce pays, où les habitants se répartissent en deux zones, reste constante dans le temps.

Chaque année, 0,5 % de citadins quittent la ville pour la campagne alors que 1,3 % de ruraux migrent vers la ville. Pour tout entier naturel n, on note c_n le nombre de citadins, en millions, en $2019+n$ et r_n le nombre de ruraux. On a donc $c_0 = 65$ et $r_0 = 15$

On peut alors traduire ces informations par le système suivant : $\begin{cases} c_{n+1} = 0{,}995c_n + 0{,}013r_n \\ r_{n+1} = 0{,}005c_n + 0{,}987r_n \end{cases}$

A l'aide des suites de matrices colonnes, on peut alors écrire :

$\begin{pmatrix} c_{n+1} \\ r_{n+1} \end{pmatrix} = \begin{pmatrix} 0{,}995 & 0{,}013 \\ 0{,}005 & 0{,}987 \end{pmatrix} \times \begin{pmatrix} c_n \\ r_n \end{pmatrix}$ soit en posant $U_n = \begin{pmatrix} c_n \\ r_n \end{pmatrix}$ et $A = \begin{pmatrix} 0{,}995 & 0{,}013 \\ 0{,}005 & 0{,}987 \end{pmatrix}$,
$\forall n \in \mathbb{N}, U_{n+1} = A \times U_n$ et $U_0 = \begin{pmatrix} 65 \\ 15 \end{pmatrix}$

On reconnait l'expression d'une suite géométrique bien connue de tous…

Propriété :

Soit A une matrice carrée d'ordre p ($p \in \mathbb{N}$) et (U_n) une suite de matrices colonnes à p lignes telles que $\forall n \in \mathbb{N}, U_{n+1} = A \times U_n$. Alors, $\forall n \in \mathbb{N}, U_n = A^n \times U_0$

Programmation en Python :

A l'aide d'une boucle fermée, on peut déterminer un terme de la suite de matrice à la demande.

```
> population(0)
array([[65],
       [15]])
> population(10)
array([[63.80040143],
       [16.19959857]])
```

```python
from numpy import *
from numpy.linalg import *

def population(n):
  c,r = 65,15
  for i in range(n):
    c,r = 0.995*c + 0.013*r,0.005*c + 0.987*r
  U = array([[c],[r]])
  return U
```

Propriété :

Soit A une matrice carrée d'ordre p ($p \in \mathbb{N}$), B une matrice colonne à p lignes et (U_n) une suite de matrices colonnes à p lignes telles que $\forall n \in \mathbb{N}, U_{n+1} = A \times U_n + B$.
Si (U_n) converge, alors sa limite U est une matrice colonne vérifiant $U = A \times U + B$.
La matrice U est appelée l'état stable de la suite (U_n).

Remarque :

On reconnait l'expression d'une suite arithmético-géométrique étudiée en première.

Exercice :

On donne la suite de matrices colonnes (U_n) définie $\forall n \in \mathbb{N}$, par $U_{n+1} = A \times U_n + B$.
On a $A = \begin{pmatrix} 0{,}5 & 0 \\ -0{,}5 & 0{,}5 \end{pmatrix}$ et $B = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $U_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

1. Déterminer l'état stable noté U
2. Montrer que la suite définie par $V_n = U_n - U$ est une suite géométrique.
3. Donner l'expression de V_n en fonction de n.
4. On admet que, $\forall n \in \mathbb{N}, A^n = \begin{pmatrix} 0{,}5^n & 0 \\ -n0{,}5^n & 0{,}5^n \end{pmatrix}$. Donner alors l'expression de U_n
5. Déterminer la limite de (U_n).

Solution :

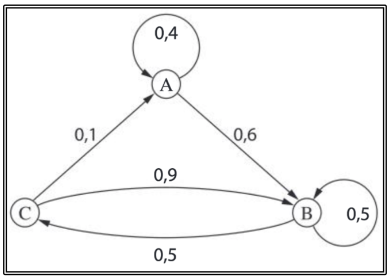
1. $U = A \times U + B \Leftrightarrow U - A \times U = B \Leftrightarrow (I_2 - A)U = B \Leftrightarrow U = (I_2 - A)^{-1}B$
   En utilisant la calculatrice, on détermine l'inverse puis on obtient $U = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$
2. $V_{n+1} = U_{n+1} - U = A \times U_n + B - U = A \times U_n + B - (A \times U + B)$
   On a : $V_{n+1} = A(U_n - U)$ soit donc $V_{n+1} = AV_n$. (V_n) est géométrique de raison A.
3. On a alors $V_n = A^n V_0$. On évalue $V_0 = \begin{pmatrix} -3 \\ 4 \end{pmatrix}$ soit $V_n = \begin{pmatrix} 0{,}5^n & 0 \\ -n0{,}5^n & 0{,}5^n \end{pmatrix} \times \begin{pmatrix} -3 \\ 4 \end{pmatrix}$
4. On inverse alors l'égalité pour obtenir : $\forall n \in \mathbb{N}, U_n = V_n + U$
   $\forall n \in \mathbb{N}, U_n = \begin{pmatrix} -3 \times 0{,}5^n + 4 \\ 3n \times 0{,}5^n + 4 \times 0{,}5^n - 2 \end{pmatrix}$
5. On détermine la limite des deux coefficients de la matrice colonne U_n.
   Au final, on a : $\lim\limits_{n \to \infty} U_n = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$

2) Les chaînes de Markov

a) Graphe orienté pondéré

Exemple introductif :

Le graphe ci-contre est d'ordre 3. Il n'est pas simple et est orienté.
La grande nouveauté, ce sont ces nombres sur les arcs. On remarque qu'ils sont tous compris entre 0 et 1 et que la somme des arcs issues de chaque sommet vaut 1. Les sommets sont ici appelés les états.



Définition :

Un graphe est pondéré lorsque chaque arête est affectée d'un nombre réel positif appelé poids de cette arête.
Un graphe probabiliste est un graphe orienté pondéré ou tous les poids sont compris entre 0 et 1 et tel que la somme des poids des arêtes issues d'un même sommet vaut 1.

Définition :

La matrice associée à un graphe probabiliste comportant p sommets est appelée matrice de transition. C'est un élément de $\mathcal{M}_{p,p}$ où le terme de la i -ème ligne et la j-ème colonne est égale au poids de l'arc allant du sommet i au sommet j s'il existe et 0 sinon.

Exemple :

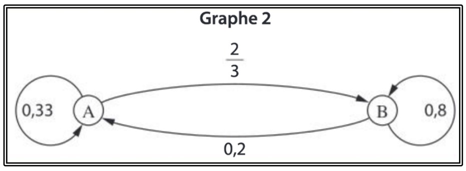
Avec le graphe probabiliste ci-dessus, on a : $A = \begin{pmatrix} 0,4 & 0,6 & 0 \\ 0 & 0,5 & 0,5 \\ 0,1 & 0,9 & 0 \end{pmatrix}$

Remarques :

- Si les sommets du graphe sont des lettres, on numérote ces sommets dans l'ordre.
- Dans la matrice de transition, le somme des termes appartenant à une même ligne vaut 1.

Contre-exemple :

Contrairement aux apparences, le graphe ci-contre n'est pas un graphe probabiliste car la somme des poids issus de A ne vaut pas 1.



b) Chaîne de Markov à deux ou trois états

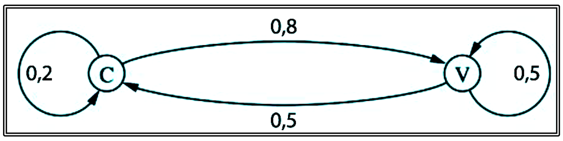
Définition :

On considère une suite de variables aléatoires (X_n) permettant de modéliser l'évolution par étapes successives d'un système aléatoire comportant différents états.
A l'étape $n = 0$, la loi de probabilité de X_0 s'appelle la distribution initiale.
A l'étape n, la loi de probabilité de X_n s'appelle la distribution après n transitions.
Lorsque, à chaque étape, la probabilité de transition d'un état à un autre ne dépend pas de n, on dit que la suite (X_n) est une chaîne de Markov.

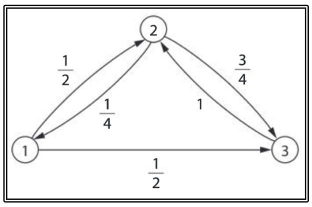
Exemples :

On donne les matrices de transition ci-dessous. Construire alors le graphe probabiliste correspondant à chaque situation.

$$A = \begin{pmatrix} 0,2 & 0,8 \\ 0,5 & 0,5 \end{pmatrix} \Longrightarrow$$



$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{3}{4} \\ 0 & 1 & 0 \end{pmatrix} \Longrightarrow$$



Remarque :

Une chaîne de Markov peut aussi se représenter par un arbre de probabilité.

3) Représentation d'une chaîne de Markov à l'aide d'une suite de matrices

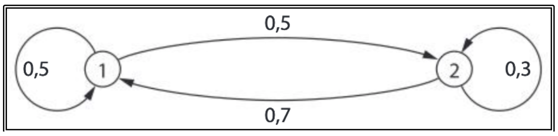
a) Modélisation à l'aide d'une suite de matrice

Propriété :

On considère une chaîne de Markov à 2 (respectivement 3) états et P la matrice de transition associée.
Soient, n, i et j trois entiers naturels tels que $n \geq 1, 1 \leq i \leq 2$ et $1 \leq j \leq 2$ (respectivement $1 \leq i \leq 3$ et $1 \leq j \leq 3$).
La probabilité de passer de l'état i à l'état j en n étapes est égale au terme de la i-ème ligne et j-ème colonne de la matrice P^n

Exemple :

On considère une marche aléatoire à deux états modélisée par le graphe probabiliste suivant et P la matrice de transition associée.



On lit alors : $P = \begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}7 & 0{,}3 \end{pmatrix}$. Pour déterminer la probabilité de passer de l'état 1 à l'état 2 en 3 étapes, on calcule avec la calculatrice $P^3 = \begin{pmatrix} 0{,}58 & 0{,}42 \\ 0{,}588 & 0{,}412 \end{pmatrix}$

On extrait alors la probabilité de 0,42 pour passer de l'état 1 à l'état 2 en 3 étapes.

Remarque :

Il est important de noter que $0{,}58 + 0{,}42 = 1$ et $0{,}588 + 0{,}412 = 1$.

b) Étude asymptotique

Définition :

On considère une chaîne de Markov à 2 (respectivement 3) états et P la matrice de transition associée.
Soit π_n la matrice ligne à 2 (respectivement 3) colonnes dont le terme de la j-ème colonne est la probabilité qu'à l'étape n, la variable aléatoire X_n soit égale à j.
Autrement dit, $\pi_n = \begin{pmatrix} P(X=1) & P(X=2) \end{pmatrix}$.
Avec 3 états, on aurait : $\pi_n = \begin{pmatrix} P(X=1) & P(X=2) & P(X=3) \end{pmatrix}$

Remarque :

- π_0 représente la distribution initiale.
- π_n représente la distribution après n transitions.

Propriétés :

$\forall n \geq 1, \pi_{n+1} = \pi_n \times P$ et $\pi_n = \pi_0 \times P^n$
S'il existe un entier n tel que P^n ne contient aucun 0, alors la suite (π_n) converge vers la matrice π vérifiant $\pi = \pi \times P$ et cette limite ne dépend pas de π_0
On dit que π représente la distribution invariante du système.

Remarque :

Dire que la matrice P^n ne contient pas de 0 signifie qu'en n étapes, on peut passer de n'importe quel état à n'importe quel autre.

c) Distribution invariante

Exemple :

On considère une chaîne de Markov dont la matrice de transition est : $P = \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}9 & 0{,}1 \end{pmatrix}$

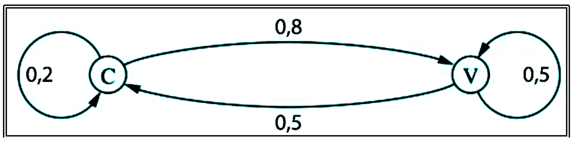
Alors, : $\pi = (0{,}6 \quad 0{,}4)$ est une distribution invariante.

En effet : $(0{,}6 \quad 0{,}4) \times \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}9 & 0{,}1 \end{pmatrix} = (0{,}6 \times 0{,}4 + 0{,}4 \times 0{,}9 \quad 0{,}6 \times 0{,}6 + 0{,}4 \times 0{,}1)$

On obtient bien que $\pi \times P = \pi$

Exercice :

On a programmé un ordinateur pour qu'il affiche successivement des lettres qui sont soit des consonnes (C) soit des voyelles (V) selon le graphe probabiliste ci-dessous.



1) On suppose que la première lettre est une consonne. Calculer la probabilité que la cinquième lettre soit une consonne.
2) Déterminer la distribution invariante de ce système. Interpréter le résultat.

Solution :

1) La matrice de transition de cette chaîne de Markov est $P = \begin{pmatrix} 0{,}2 & 0{,}8 \\ 0{,}5 & 0{,}5 \end{pmatrix}$. Puisque la première lettre est une consonne, on a : $\pi_1 = (1 \quad 0)$.

On a donc : $\pi_5 = \pi_1 \times P^4$ soit $\pi_5 = (1 \quad 0) \times \begin{pmatrix} 0{,}2 & 0{,}8 \\ 0{,}5 & 0{,}5 \end{pmatrix}^4$.

On obtient $\pi_5 = (0{,}3896 \quad 0{,}6104)$.
La probabilité que la cinquième lettre soit une consonne est 0,3896.

2) On résout $\pi = \pi \times P$ soit alors $(x \quad y) = (x \quad y) \times \begin{pmatrix} 0{,}2 & 0{,}8 \\ 0{,}5 & 0{,}5 \end{pmatrix}$.En écrivant sous forme d'un système, on a :$\begin{cases} -0{,}8x + 0{,}5y = 0 \\ 0{,}8x - 0{,}5y = 0 \end{cases}$. Les deux équations sont équivalentes.

Avec la contrainte $x + y = 1$, on doit résoudre$\begin{cases} x + y = 1 \\ 0{,}8x - 0{,}5y = 0 \end{cases}$.

Avec une substitution simple, on a : $\begin{cases} x = 1 - y \\ 0{,}8(1 - y) - 0{,}5y = 0 \end{cases} \Leftrightarrow \begin{cases} x = \frac{5}{13} \\ y = \frac{8}{13} \end{cases}$

Puisque la matrice P ne contient pas de 0, les distributions de probabilités convergent vers cette distribution invariante. Ainsi, à l'infini, la probabilité d'obtenir une consonne vaut $\frac{5}{13}$ et une voyelle $\frac{8}{13}$.