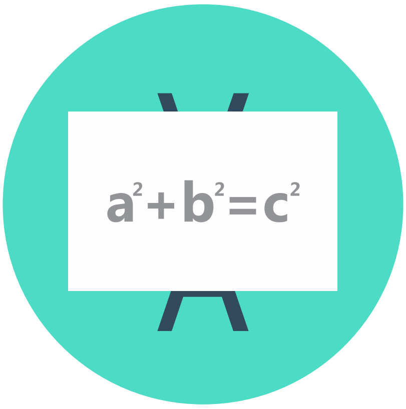


Chapitre XVI – Chaînes de Markov (Maths expertes)

Bacomathiques — https://bacomathiqu.es

TABLE DES MATIÈRES

I – Graphe pondéré et graphe probabiliste

1. Définition

À RETENIR

Graphe pondéré

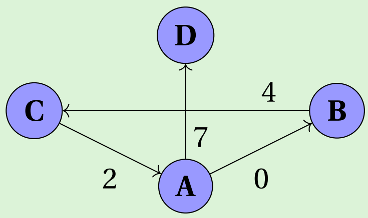
Un graphe est dit **pondéré** si chacune de ses arêtes est affecté d'un nombre positif (ou nul) que l'on appelle **poids**.

Le poids d'une chaîne (ou d'un chemin) est la somme des poids de ses arêtes.

À LIRE

Exemple

On considère le graphe orienté et pondéré suivant :



On a :
- Le poids de l'arête $A-B$ vaut 0.
- Le poids du chemin $A-B-C-A-D$ vaut $0+4+2+7=13$.

À RETENIR

Graphe probabiliste

On appelle **graphe probabiliste** un graphe orienté et pondéré tel que :
- Pour chaque sommet, la somme des poids des arcs issus de ce sommet vaut 1.
- Il y a au plus 1 arrête orientée reliant chaque sommet.

Il peut être utile de faire l'analogie entre les graphes probabilistes et les arbres de probabilité vus en classe de Première.

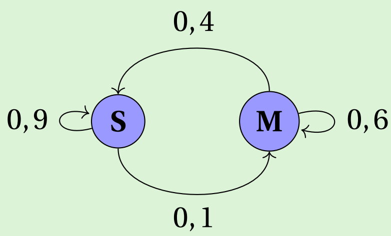
À LIRE

Exemple

Faisons un exemple concret. On souhaite étudier l'évolution d'une maladie chez un certain individu. À un jour donné, cet individu est soit malade (que l'on note M), soit soigné (que l'on note S). On suppose que pour cette maladie :

- La probabilité qu'une personne malade guérisse le lendemain est $0,4$.
- La probabilité qu'une personne saine tombe malade le lendemain est $0,1$.

Le graphe probabiliste modélisant cette situation est le graphe G suivant :



On remarque que la somme des poids des arêtes issues du sommet S vaut $0,9+0,1=1$ (idem pour M qui vaut $0,6+0,4=1$).

2. Matrice de transition

À RETENIR

Définition

Soit G un graphe probabiliste d'ordre n. On appelle **matrice de transition** du graphe G, la matrice carrée d'ordre n dont le coefficient à la ligne i et à la colonne j est égal au poids de l'arête reliant le sommet i au sommet j.

Une telle matrice est qualifiée de **stochastique** car la somme des coefficients de chacune de ses lignes vaut 1.

À LIRE

Exemple

Dans l'exemple précédent (en supposant que S est le 1$^{\text{er}}$ sommet et que M est le 2$^{\text{e}}$) la matrice de transition du graphe G est $\begin{pmatrix} 0,9 & 0,1 \\ 0,4 & 0,6 \end{pmatrix}$.

Attention cependant à ne pas confondre matrice de transition et matrice d'adjacence.

II – Chaînes de Markov

1. Qu'est-ce qu'une chaîne de Markov ?

Il vous est fortement conseillé de relire (et de maîtriser) le cours sur les variables aléatoires avant d'aborder cette section. De plus, sachez que cette partie est sans doute la plus difficile du programme de Terminale. Mais ne vous découragez pas car elle reste parfaitement accessible !

À RETENIR

Définition

Soit (X_n) une suite de variables aléatoires discrètes définies sur un même univers Ω et à valeurs dans un ensemble E. On dit que (X_n) définit une **chaîne de Markov** sur E si pour tout $n \in \mathbb{N}$ et tout $x_0, x_1, x_2, \dots, x_n \in E$, l'événement $(X_n = x_n)$ ne dépend que de l'événement antérieur $(X_{n-1} = x_{n-1})$ (et pas des précédents) ; autrement dit, si $P_{(X_{n-1}=x_{n-1})\cap\dots\cap(X_0=x_0)}(X_n = x_n) = P_{(X_{n-1}=x_{n-1})}(X_n = x_n)$.

De plus, l'ensemble E est appelé **espace des états** de la chaîne de Markov.

En français, cela signifie que si X_n représente l'état d'un système à un temps n, alors l'état suivant X_{n+1} ne dépend que de l'état au temps n et pas des états précédents. De plus, notez bien que nous n'avons pas fait d'hypothèse sur le cardinal de E (qui peut donc être de cardinal $m \in \mathbb{N}$).

En classe de Terminale, nous nous limiterons principalement au cas où E possède 2 voire 3 éléments, mais nous allons quand-même voir très bientôt un exemple de chaîne de Markov à 12 états.

À LIRE

Variable aléatoire discrète

Une variable aléatoire X définie sur un univers Ω est dite **discrète** si $X(\Omega)$ est un ensemble dénombrable.

À RETENIR

Chaîne de Markov homogène

Soit (X_n) une chaîne de Markov dont on note E l'espace des états. Alors (X_n) est dite **homogène** si pour tout $n \in \mathbb{N}$ et pour tout $x, y \in E$, la probabilité $P_{(X_n=x)}(X_{n+1} = y)$ est indépendante de n.

En termes mathématiques, cela signifie que pour tout $n \in \mathbb{N}$ et pour tout $x, y \in E$, $P_{(X_n=x)}(X_{n+1} = y) = P_{(X_0=x)}(X_1 = y)$.

À LIRE

Exemple

Eliott fait la collection des vignettes des 11 joueurs titulaires de l'Équipe de France de football qu'il trouve dans des paquets de céréales. À chaque fois qu'il achète un paquet, il a donc une probabilité de $\frac{1}{11}$ de tomber sur le k-ième joueur (pour tout k compris entre 1 et 11).

Si on note par X_n le nombre de vignettes différentes dans la collection d'Eliott après qu'il eut ouvert n paquets de céréales, alors (X_n) est une chaîne de Markov homogène (commençant par $X_0 = 0$). En effet, pour tout $k \in \{0, 1, \dots, 11\}$, on a que l'événement $(X_{n+1} = k)$ ne dépend que de X_n :

$$P_A(X_{n+1} = k) = \begin{cases} \frac{k}{11} \text{ si } A \text{ est l'événement } (X_n = k) \\ 1 - \frac{k-1}{11} \text{ si } A \text{ est l'événement } (X_n = k-1) \text{ et que } k \geq 1 \\ 0 \text{ sinon} \end{cases}$$

Pour détailler un peu plus :

— Si on a $(X_n = k)$ (i.e. on a déjà tiré k joueurs différents), alors la probabilité d'avoir $(X_{n+1} = k)$ est égale à la probabilité de ne pas tirer de nouveau joueur (qui est $\frac{k}{11}$). Cela inclut également le cas où $k = 11$.

— Si on a $(X_n = k-1)$ (i.e. on a déjà tiré $k-1$ joueurs différents), alors la probabilité d'avoir $(X_{n+1} = k)$ est égale à la probabilité de tirer un nouveau joueur (qui est $1 - \frac{k-1}{11}$).

— Sinon, comme on ne peut pas tirer plus d'un nouveau joueur d'un coup ou en enlever de la collection, la probabilité d'avoir $(X_{n+1} = k)$ est égale à 0.

— Notons de plus que (X_n) est homogène car le calcul de $P(X_{n+1} = k)$ est indépendant de n (mais reste dépendant de X_n, attention).

De plus, l'espace des états E est $\{0, 1, \dots, 11\}$.

Cet exemple est très connu et porte un nom : il s'agit du **problème du collectionneur de vignettes**. Pour votre culture, sachez qu'en moyenne, il faudra ouvrir environ $n \ln(n)$ paquets de céréales pour compléter une collection de n vignettes.

2. Matrice et graphe associés à une chaîne de Markov

À RETENIR

Matrice de transition

Soit (X_n) une chaîne de Markov homogène dont on note $E = \{x_1, x_2, \dots, x_m\}$ l'espace des états. La **matrice de transition** de (X_n) est la matrice carrée d'ordre m dont le coefficient situé à la i-ième ligne et à la j-ième colonne est égal à $p_{i,j} = P_{(X_n = x_i)}(X_{n+1} = x_j)$.

À LIRE

Comme cette probabilité est indépendante de n, on peut tout à fait prendre $n=0$ dans la définition. On a alors $p_{i,j} = P_{(X_0=x_i)}(X_1 = x_j)$.

À RETENIR

Graphe associé à une chaîne de Markov

Soit (X_n) une chaîne de Markov homogène dont on note $E = \{x_1, x_2, \dots, x_m\}$ l'espace des états. On associe à cette chaîne de Markov un graphe probabiliste G d'ordre m dont les sommets sont les états x_i et dont les arêtes $x_i - x_j$ sont pondérées par les poids $p_{i,j} = P_{(X_n=x_i)}(X_{n+1} = x_j)$.

La matrice de transition de (X_n) est égale à la matrice de transition du graphe probabiliste G : il s'agit donc aussi d'une matrice stochastique.

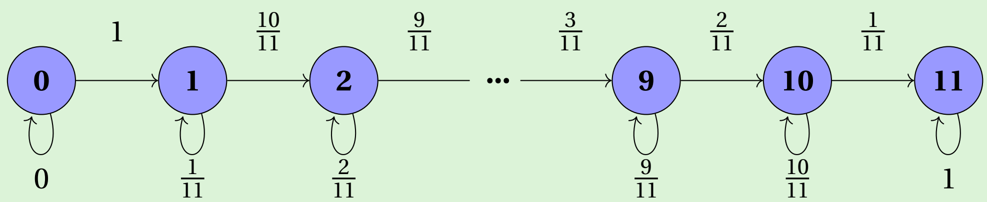
À LIRE

Exemple

Reprenons l'exemple précédent. Alors la matrice de transition associée à (X_n) est la matrice $M \in \mathscr{M}_{12}(\mathbb{R})$:

$$M = \begin{pmatrix} 0 & 1 & \dots & 0 & 0 \\ 0 & \frac{1}{11} & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \frac{10}{11} & \frac{1}{11} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

Et le graphe associé à (X_n) est le graphe probabiliste d'ordre 12 :



3. Distributions dans une chaîne de Markov

À RETENIR

Proposition

Soit (X_n) une chaîne de Markov homogène dont on note $E = \{x_1, x_2, \dots, x_m\}$ l'espace des états. On pose $p_{i,j}^{(k)} = P_{(X_0=x_i)}(X_k = x_j)$ pour tout $k \in \mathbb{N}^*$ (qui représente la probabilité que la chaîne de Markov (X_n) passe de l'état x_i à l'état x_j en k étapes). On a :

$$p_{i,j}^{(k)} = \sum_{q=1}^{m} p_{i,q}^{(k-1)} \times p_{q,j}^{(1)} = p_{i,1}^{(k-1)} \times p_{1,j}^{(1)} + p_{i,2}^{(k-1)} \times p_{2,j}^{(1)} + \cdots + p_{i,m}^{(k-1)} \times p_{m,j}^{(1)}.$$

De plus, comme (X_n) est homogène, $p_{i,j}^{(k)} = p_{i,j}^{(n+k)}$ pour tout $n \in \mathbb{N}$.

DÉMONSTRATION

Proposition

$p_{i,j}^{(k)} = P_{(X_0=x_i)}(X_k = x_j)$

$= \sum_{q=1}^{m} P_{(X_0=x_i)}((X_k = x_j) \cap (X_{k-1} = x_q))$

$= \sum_{q=1}^{m} P_{(X_{k-1}=x_q)\cap(x_0=x_i)}(X_k = x_j) P_{(X_0=x_i)}(X_{k-1} = x_q)$ (par la formule des probabilités totales)

$= \sum_{q=1}^{m} P_{(X_{k-1}=x_q)}(X_k = x_j) P_{(X_0=x_i)}(X_{k-1} = x_q)$

$= \sum_{q=1}^{m} P_{(X_0=x_q)}(X_1 = x_j) P_{(X_0=x_i)}(X_{k-1} = x_q)$ (par homogénéité)

$= \sum_{q=1}^{m} p_{j,q}^{(1)} \times p_{i,q}^{(k-1)}$.

Cette formule semble un petit peu compliquée à interpréter. Elle signifie simplement que la probabilité que la chaîne de Markov (X_n) passe de l'état x_i à l'état x_j en k étapes est égale à la probabilité qu'elle passe de l'état e_i à e_q en une étape, puis de passer de e_q à e_j en $k-1$ étapes. Heureusement, il est possible de la simplifier grandement à l'aide des matrices de transition.

À RETENIR

Lien avec la matrice de transition

En reprenant les notations précédentes et en notant M la matrice de transition de (X_n), alors $p_{i,j}^{(k)}$ est le coefficient à la ligne i et à la colonne j de la matrice M^k.

Enfin, donnons la définition centrale de cette section.

À RETENIR

Définition

Soit (X_n) une chaîne de Markov homogène dont on note $E = \{x_1, x_2, \dots, x_m\}$ l'espace des états. On appelle **suite des distributions** de (X_n) la suite de matrices (π_n), définie pour tout $n \in \mathbb{N}$ par $\pi_n = \begin{pmatrix} P(X_n = x_1) & P(X_n = x_2) & \dots & P(X_n = e_m) \end{pmatrix}$.

π_n est donc une matrice ligne d'ordre m et est appelée **distribution au temps** n.

π_0 (la distribution au temps 0) est appelée **distribution initiale**.

Une propriété très sympathique des distributions, est que l'on dispose d'une relation de récurrence permettant de calculer facilement la distribution à un temps n donné.

À RETENIR

Relation entre π_{n+1} et π_n

En reprenant les notations de la définition précédente et en notant M la matrice de transition de (X_n), alors la suite (π_n) vérifie une relation de récurrence donnée pour tout $n \in \mathbb{N}$ par $\pi_{n+1} = \pi_n M$.

On en déduit que pour tout $n \in \mathbb{N}$, $\pi_n = \pi_0 M^n$.

DÉMONSTRATION

Relation entre π_{n+1} et π_n

Soit $n \in \mathbb{N}$. Les événements $(X_n = x_1), (X_n = x_2), \dots, (X_n = x_m)$ partitionnent (recouvrent) notre univers, donc par la formule des probabilités totales appliquée à notre système complet d'événements et à $(X_{n+1} = x_j)$:

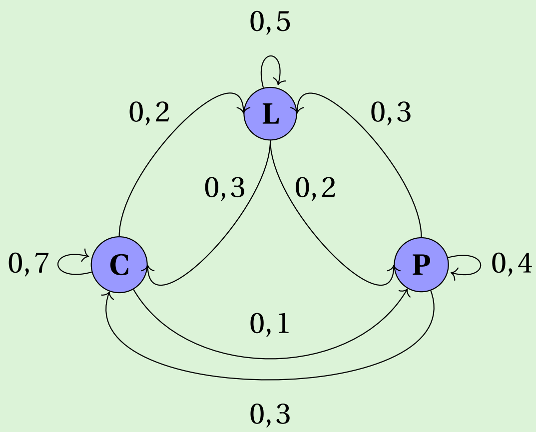
$$\begin{aligned}
&P(X_{n+1} = x_j) = P((X_{n+1} = x_j) \cap (X_n = x_1)) + \dots + P((X_{n+1} = x_j) \cap (X_n = x_m)) \\
&= P_{(X_n = x_1)}(X_{n+1} = x_j) \times P(X_n = x_1) + \dots + P_{(X_n = x_m)}(X_{n+1} = x_j) \times (X_n = x_m) \\
&= \pi_n M
\end{aligned}$$

Et la formule $\pi_n = \pi_0 M^n$ se déduit de la formule d'une suite géométrique (où M serait "la raison" et π_0 le premier terme).

À LIRE

Exemple

Intéressons-nous à l'alimentation d'un chat durant la journée. Il dispose de trois gamelles différentes L, C et P dans lesquelles se trouvent respectivement du lait, des croquettes et de la pâté. On suppose que le chat a commencé sa journée par du lait et que toutes les heures, il se dirige vers une des gamelles suivant le graphe probabiliste ci-dessous :



On note par X_n la variable aléatoire qui donne la gamelle qu'a choisi le chat à la n-ième heure. On a donc que (X_n) est une chaîne de Markov homogène dont l'espace des états est $E = \{L; C; P\}$. Si on note (π_n) la suite des distributions de (X_n), on a alors que $\pi_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$.

Soit M la matrice de transition M de (X_n). Calculons quelques puissances de M :

— $M = \begin{pmatrix} 0,5 & 0,3 & 0,2 \\ 0,2 & 0,7 & 0,1 \\ 0,3 & 0,3 & 0,4 \end{pmatrix}$

— $M^2 = \begin{pmatrix} 0,37 & 0,42 & 0,21 \\ 0,27 & 0,58 & 0,15 \\ 0,33 & 0,42 & 0,25 \end{pmatrix}$

— $M^3 = \begin{pmatrix} 0,332 & 0,468 & 0,2 \\ 0,296 & 0,532 & 0,172 \\ 0,324 & 0,468 & 0,208 \end{pmatrix}$

Ainsi :

— $\pi_1 = \pi_0 M = \begin{pmatrix} 0,5 & 0,3 & 0,2 \end{pmatrix}$

— $\pi_2 = \pi_0 M^2 = \begin{pmatrix} 0,37 & 0,42 & 0,21 \end{pmatrix}$

— $\pi_3 = \pi_0 M^3 = \begin{pmatrix} 0,332 & 0,468 & 0,2 \end{pmatrix}$

Et par exemple $p_{1,2}^{(3)} = 0,468$: la probabilité que le chat passe à sa gamelle de croquettes 3 heures après le lait est d'environ 47 %.

4. Distribution invariante

À RETENIR

Définition

Soit (X_n) une chaîne de Markov homogène de matrice de transition M. Une distribution π est **invariante** si les deux conditions suivantes sont respectées :

- $\pi M = \pi$ (donc si π est une distribution à un temps n, on a $\pi = \pi_n$ et cette condition se résume à avoir $\pi_n = \pi_n M = \pi_{n+1}$).
- La somme des coefficients de π vaut 1.

À RETENIR

Existence et unicité de la distribution invariante au temps n

Soit (X_n) une chaîne de Markov homogène de matrice de transition M.

Si M ne possède aucun coefficient non nul autre que sur sa diagonale, alors (X_n) admet une unique distribution invariante π.

À RETENIR

Convergence de la distribution

Soit (X_n) une chaîne de Markov homogène dont on note (π_n) la suite des distributions.

- Si (π_n) est une suite de matrices convergente, alors elle converge vers une distribution invariante π.
- Si le cardinal de l'ensemble des états de (X_n) est 2, alors (π_n) est convergente (et converge vers la distribution invariante π).

À LIRE

Exemple

Reprenons l'exemple précédent et voyons si (X_n) admet une distribution invariante.

Remarquons tout d'abord que la matrice de transition M ne possède aucun coefficient non nul. Donc (X_n) admet une unique distribution invariante π.

Posons donc $\pi = \begin{pmatrix} x & y & z \end{pmatrix}$ et déterminons x, y et z :

On doit avoir $\pi M = \pi$

$$\iff \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 0,5 & 0,3 & 0,2 \\ 0,2 & 0,7 & 0,1 \\ 0,3 & 0,3 & 0,4 \end{pmatrix} = \begin{pmatrix} x & y & z \end{pmatrix}$$

$$\iff \begin{pmatrix} 0,5x + 0,2y + 0,3z & 0,3x + 0,7y + 0,3z & 0,2x + 0,1y + 0,4z \end{pmatrix} = \begin{pmatrix} x & y & z \end{pmatrix}$$

$$\iff \begin{cases} \frac{1}{2}x + \frac{1}{5}y + \frac{3}{10}z = x \\ \frac{3}{10}x + \frac{7}{10}y + \frac{3}{10}z = y \\ \frac{1}{5}x + \frac{1}{10}y + \frac{2}{5}z = z \end{cases} \quad \text{(en passant en écriture fractionnaire)}$$

$$\iff \begin{cases} -\frac{1}{2}x + \frac{1}{5}y + \frac{3}{10}z = 0 \\ \frac{3}{10}x - \frac{3}{10}y + \frac{3}{10}z = 0 \\ \frac{1}{5}x + \frac{1}{10}y - \frac{3}{5}z = 0 \end{cases}$$

$$\iff \begin{cases} x = \frac{2}{5}y + \frac{3}{5}z \\ -\frac{9}{50}y + \frac{12}{25}z = 0 \\ \frac{9}{50}y - \frac{12}{25}z = 0 \end{cases}$$

$$\iff \begin{cases} x = \frac{2}{5} \times \frac{8}{3}z + \frac{3}{5}z = \frac{5}{3}z \\ y = \frac{50}{9} \times \frac{12}{25}z = \frac{8}{3}z \end{cases}$$

Donc π est de la forme $\pi = \begin{pmatrix} \frac{5}{3}z & \frac{8}{3}z & z \end{pmatrix}$. De plus, la somme des coefficients de π doit faire 1, donc :

$\frac{5}{3}z + \frac{8}{3}z + z = 1 \iff \frac{16}{3}z = 1 \iff z = \frac{3}{16}$.

Donc l'unique distribution invariante π de (X_n) est
$\pi = \begin{pmatrix} \frac{5}{16} & \frac{1}{2} & \frac{3}{16} \end{pmatrix} = \begin{pmatrix} 0,3125 & 0,5 & 0,1875 \end{pmatrix}$.