Chapitre 11 - graphes probabilistes - chaines de Markov

Objectifs d'apprentissage - A la fin de ce chapitre, je sais :
- représenter un graphe probabiliste à partir des données d'un problème ☐
- déterminer la matrice de transition d'une chaine de Markov ☐
- utiliser la formule des probabilités totales pour démontrer la relation entre deux états consécutifs ☐
- détermimer les éventuels états stables d'une chaine de Markov ☐
- ... toujours diagonaliser des matrices afin d'étudier leurs puissances ☐

Dans tout le chapitre, $(X_n)_{n\in\mathbb{N}}$ désigne une suite de variables aléatoires à valeurs dans $\mathbb{N}$, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$. Pour tout $n \in \mathbb{N}, X_n(\Omega)$ est fini, et plus précisément $\exists r \in \mathbb{N}^*, \forall n \in \mathbb{N}, X_n(\Omega) \subset [\![1, r]\!]$ (pour tout n, X_n est à valeurs dans un même ensemble fini d'entiers).

1 Chaines de Markov, états, matrice de transition

Définitions : $(X_n)_{n\in\mathbb{N}}$ est appelée **chaine de Markov**, si la loi de X_{n+1} sachant $X_0, \dots X_n$ est celle de X_{n+1} sachant X_n, ce qui se traduit par :
pour tout $n \in \mathbb{N}$, et tous entiers $i_0, \dots i_{n+1}$,

$$P_{[X_n=i_n]\cap[X_{n-1}=i_{n-1}]\cap\ldots\cap[X_0=i_0]}(X_{n+1} = i_{n+1}) = P_{[X_n=i_n]}(X_{n+1} = i_{n+1})$$

sous réserve que $P([X_n = i_n] \cap [X_{n-1} = i_{n-1}] \cap \ldots \cap [X_0 = i_0]) \neq 0$

on appelle **état à l'instant** n ou $n^{\text{ème}}$ **état** de la chaine de Markov, noté V_n la matrice ligne $V_n = \begin{pmatrix} P(X_n = 1) & \ldots & P(X_n = r) \end{pmatrix}$ (où $r \in \mathbb{N}^*$)

Remarques :
- en général, cette suite $(X_n)_{n\in\mathbb{N}}$ sera donnée par un processus aléatoire concret (cf. exemples et exercices). Dans ces exemples concrets, n correspond à un moment temporel.
- pour une chaine de Markov, l'état à l'instant $n + 1$ (les valeurs possibles de X_{n+1}) ne dépend que de l'état à l'instant n
- selon la situation, on peut avoir $(X_n)_{n\in\mathbb{N}}$ (le problème commence à l'instant 0) ou $(X_n)_{n\in\mathbb{N}^*}$ (le problème commence à l'instant 1)

Définitions :
pour $i, j \in [\![1, r]\!]^2$, la probabilité $P_{X_n=i}(X_{n+1} = j)$ est appelée **probabilité de transition** de l'état i à l'état j

la matrice $M = (m_{i,j})$ avec $m_{i,j} = P_{X_n=i}(X_{n+1} = j)$, est appelée **matrice de transition**

lorsque la matrice M ne dépend pas de n, on parle de chaine de Markov **homogène**

Remarques :
- pour nous, les chaines de Markov seront (presque) toujours homogènes ;
- pour des raisons pratiques, en particulier en vue d'une programmation Python, il se peut qu'on prenne $X_n(\Omega) \subset \{0, 1, \dots r\}$ et donc d'indicer les lignes et les colonnes de la matrice de transition à partir de 0

Propriétés :

- d'après la formule des probabilités totales ,

$$\boxed{\forall n \in \mathbb{N}, \quad \forall j \in [\![1, r]\!], \quad P(X_{n+1} = j) = \sum_{i=1}^{r} P(X_n = i) P_{[X_n = i]}(X_{n+1} = j)}$$

- matriciellement cela s'écrit $\boxed{\forall n \in \mathbb{N}, \quad V_{n+1} = V_n M}$

Remarques :

- le système complet d'événements est $(X_n = i)_{i \in [\![1,r]\!]} = ((X_n = 1), (X_n = 2), \dots, (X_n = r))$ (à un instant n, ce sont toutes les valeurs possibles pour X_n)
- on écrit aussi $P(X_{n+1} = j) = \sum_{i=1}^{r} P(X_n = i) m_{i,j}$ où $M = (m_{i,j})_{(i,j) \in [\![1,r]\!]^2}$

Définition :

pour une chaine de Markov homogène à valeurs dans $[\![1, r]\!]$, le graphe orienté pondéré à r sommets, numérotés de 1 à r, contenant une arête pondérée par $m_{i,j}$ entre i et j si $m_{i,j} > 0$ (où $m_{i,j} = P_{[X_n = i]}(X_{n+1} = j)$) est appelé **graphe probabiliste**

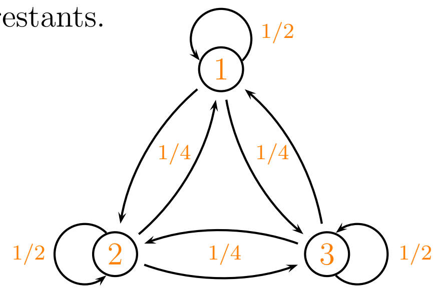
Exemple : une puce se déplace aléatoirement sur trois sommets, à chaque déplacement,

- soit elle reste sur le même sommet avec une probabilité $\frac{1}{2}$
- soit elle change de sommet avec équiprobabilité entre les sommets restants.

alors la matrice de transition est :

$$M = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

et le graphe probabiliste :



⚠ La matrice de transition peut être déterminée à l'aide du graphe ou des probabilités conditionnelles (de transition) mais c'est la formule des probabilités totales qui démontre le lien $V_{n+1} = V_n M$

Quelques résultats sur les matrices stochastiques

Les résultats de cette section, ne sont pas au programme, mais bons à connaitre.

Une matrice carrée $A = (a_{i,j})$ d'ordre r sera dite **stochastique** si :

- tous les coefficients sont positifs : $\forall (i,j) \in [\![1, r]\!]^2, \quad a_{i,j} \geqslant 0$
- la somme des éléments de chaque ligne vaut 1 : $\forall i \in [\![1, r]\!], \quad \sum_{j=1}^{r} a_{i,j} = 1$

Dans ce chapitre (chaines de Markov), les matrices de transition sont toujours des matrices stochastiques car $\forall i \in [\![1, r]\!]$,

$$\sum_{j=1}^{r} P_{[X_n=i]}(X_{n+1} = j) = P_{[X_n=i]}(X_{n+1} = 1) + P_{[X_n=i]}(X_{n+1} = 2) + \dots + P_{[X_n=i]}(X_{n+1} = r) = 1$$

i.e. pour chaque valeur i de X_n (soit la ligne de i de la matrice), l'ensemble des valeurs pour l'instant suivant vaut 1 (i.e. la somme des colonnes).

Des résultats classiques sur les matrices stochastiques (les deux premiers sont faciles à retrouver).
Si A est une matrice stochastique d'ordre r, alors :

- $1 \in \mathrm{Sp}(A)$
- le vecteur colonne ${}^t(1, \ldots 1)$ est un vecteur propre de A associé à la valeur propre 1
- $\mathrm{Sp}(A) \subset [-1, 1]$. Autrement dit : $\lambda \in \mathrm{Sp}(A) \Longrightarrow |\lambda| \leqslant 1$

2 Etat stable

On considère une chaine de Markov homogène $(X_n)_{n\in\mathbb{N}}$, telle que $X_n(\Omega) = [\![1, r]\!]$
on note toujours $M = (m_{i,j})$ la matrice de transition associée
et $V_n = \begin{pmatrix} P(X_n = 1) & \ldots & P(X_n = r) \end{pmatrix}$ qui vérifie donc $\forall n \in \mathbb{N}, \quad V_{n+1} = V_n M$

Définition :
un état V (i.e. une matrice ligne) $V = \begin{pmatrix} v_1 & \ldots & v_r \end{pmatrix}$ est appelé **état stable** s'il vérifie :

(i) $\forall i \in [\![1, r]\!]$, $v_i \geqslant 0$ et $\displaystyle\sum_{i=1}^{r} v_i = 1$ (i.e. V définit une loi de probabilité)

(ii) $V = VM$

Remarques :

- un état stable est donc d'un état théorique où la loi de resterait constante (ce qui ne veut pas dire qu'il n'y a pas de déplacement entre les sommets), on parle également de loi de probabilité invariante ;
- V est un état stable $\Leftrightarrow {}^tV$ est un vecteur propre de la matrice de transition transposée ${}^tM$ car $V = VM \Leftrightarrow {}^tV = {}^t(VM) = {}^tM\,{}^tV$
- si les composantes de V_n ont une limite quand $n \to +\infty$, alors la matrice dont les coefficients sont les limites définit un état stable

3 L'étude classique d'une chaine de Markov

Les finalités de ces études, qui font l'objet de problèmes entiers sont :

- le calcul de V_n en fonction de n
- le calcul de $E(V_n)$ en fonction de n
- l'évolution de V_n ou de $E(V_n)$ en fonction de n, en particulier quand $n \to +\infty$

Avant cela, il faut :

- déterminer la matrice de transition avec les probabilités conditionnelles
- établir le lien $V_{n+1} = V_n M$ avec la formule des probabilités totales
- diagonaliser la matrice M
- montrer que $V_n = V_0 M^n$ (ou $V_n = V_1 M^{n-1}$) et expliciter M en montrant que $M^n = PD^nP^{-1}$

On récapitule avec la marche aléatoire sur les sommets d'un polygone

Un mobile se déplace sur les sommets d'un polygone régulier à r côtés numérotés $1, 2, \ldots, r$
Lorsqu'il est sur un sommet à l'étape n, il passe à un des deux sommets adjacents à l'étape $n+1$ avec équiprobabilité.
Pour $n \in \mathbb{N}^*$, on note X_n la position du mobile (état) au bout de k étapes, avec $X_0 = 1$

Pour l'exemple, on choisit $r = 3$ (triangle).

On a ainsi : $\forall i \in \{1, 2, 3\}, \forall j \in \{1, 2, 3\},\ j \neq i \quad P_{[X_n=i]}(X_{n+1} = j) = \dfrac{1}{2}$

Les probabilités $m_{i,j} = P_{[X_n=i]}(X_{n+1} = j)$ sont les probabilités de transition.

En notant $m_{i,j} = P_{X_n=i}(X_{n+1} = j)$, on a, grâce à la formule des probabilités totales, la relation :

$$\boxed{\forall j \in [\![1, 3]\!], \quad P(X_{n+1} = j) = \sum_{i=1}^{3} m_{i,j} P(X_n = i)} \tag{1}$$

La matrice $M = (m_{i,j})_{1 \leqslant i,j \leqslant 3} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ est la **matrice de transition**

La matrice de transition contient en ligne i et colonne j la probabilité de passer de l'état i à l'état j

On remarquera que la somme des éléments de chaque ligne de M vaut 1, c'est un exemple de matrice stochastique.

En notant $V_n = \begin{pmatrix} P(X_n = 1) & P(X_n = 2) & P(X_n = 3) \end{pmatrix}$, les relations de (1) se traduisent matriciellement par :

$$\boxed{\forall n \in \mathbb{N}, \quad V_{n+1} = V_n M}$$

Puis par récurrence, on obtient, $\boxed{\forall n \in \mathbb{N}, \quad V_n = V_0 M^n}$

Le schéma d'évolution de l'état n à l'état $n+1$ est représenté par un **graphe probabiliste** :