Mathématiques

Une étude des chaînes de Markov

Alain Combrouze
Professeur en classes préparatoires scientifiques MPSI, lycée St-Louis (Paris) et en classes préparatoires économiques et commerciales, Prépasup (Paris), Auteur de "Probabilités et statistiques" et de "Analyse et algorithmes" (Collection Major, PUF).

Ce petit article est avant tout un hommage à mon ami Tran Van Hiep.

J'ai profité de la proposition de Serge Le Diraison qui a eu l'idée de consacrer ce texte à sa mémoire. Ma rencontre avec Tran remonte à une quinzaine d'années maintenant, à une époque où nous commencions à enseigner en classe préparatoire aux écoles de commerce, lui à IPESUP et moi à PRÉPASUP.

Cette relation s'est trouvée renforcée par notre collaboration au premier ouvrage de la collection Major dirigée par Pascal Gauchon.

Il s'agissait de rédiger un tome d'anaIyse et d'algorithmique conforme au programme de ces classe. Au cours de cette collaboration, j'ai pu apprécier les qualités de créativité et de rigueur chez cet homme pour qui la vie n'avait pas toujours été facile.

Il avait su maintenir un cap et se forger une éthique personnelle exigeante, qui l'ont amené à donner beaucoup aux siens tout en assurant un brillant cursus en physique théorique.

Cette exigence prit même parfois le pas sur une relation d'amitié naissante et donna lieu à des tensions mémorables.

Il n'en demeure pas moins qu'elle se traduisait par une honnêteté sur le plan des idées.

J'ai eu par ailleurs l'occasion de passer quelques moments avec Tran dans le cadre de sa famille, et l'ai toujours trouvé attentif, tendre et affectueux envers sa femme et ses enfants.

C'est quelqu'un de caractère et de valeur qui s'en est allé. En dehors de ses proches, il demeurera présent dans la mémoire de ses étudiants et des collègues et amis qui l'ont approché.

Dès lors, que pouvais-je aborder d'autre comme thème dans ce texte, sinon ces processus stochastiques que l'ont dit "doués" de mémoire, et que l'on appelle chaînes de Markov ?

Ils évoluent au cours du temps en conformité avec leur histoire, et nous permettent même d'envisager une inversion du temps qui, si elle se transposait sur un plan non mathématique, ouvrirait les perspectives d'un "être mieux"...

A. C.

$\boxed{A}$

Un système évolue au cours du temps entre n états $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$.

Il est entendu que ce système ne se trouve que dans un seul des états précédents à chaque instant entier $t = k$, $k \in \mathbb{N}$.

Nous faisons alors l'hypothèse suivante :

A chaque instant entier $t = k$, il passe de l'état $\mathcal{E}_j$ à l'état $\mathcal{E}_i$ avec la probabilité $p_{i,j}$, $1 \leq i, j \leq n$.

Cela implique que la transition de l'état $\mathcal{E}_j$ à l'état $\mathcal{E}_i$ se fasse indépendamment de la valeur de $t = k$ (indépendamment de l'âge du système).

De la sorte, la distribution des probabilités avec lesquelles le système est dans les différents états possibles à l'instant $t = k+1$ ne dépend que de la distribution analogue à l'instant immédiatement antérieur $t = k$.

Un tel système a une mémoire relativement courte, puisqu'il ne se souvient que de l'instant immédiatement antérieur (passé immédiat).

Nous allons étudier l'évolution asymptotique de ces systèmes, les instants d'évolution étant $t = 0$, $t = 1, \ldots, t = k, \ldots$

- *Soit A, la matrice définie par $A = (p_{i,j})$.*

Montrons que A vérifie les deux propriétés suivantes :

(i) $\forall (i,j) \in [\![1,n]\!]^2 : p_{i,j} \geq 0$.

(ii) $\forall j \in [\![1,n]\!] : \sum_{i=1}^{n} p_{i,j} = 1$.

En effet $p_{i,j} = P(X_{k+1} = i \mid X_k = j)$, $1 \leq i, j \leq n$, est une probabilité conditionnelle, et $\sum_{i=1}^{n} p_{i,j} = \sum_{i=1}^{n} P(X_{k+1} = i \mid X_k = j) = 1$, car le système est nécessairement dans un des états $\mathcal{E}_1, \ldots, \mathcal{E}_n$ et un seul à chaque instant $k+1$ sachant qu'il était dans l'état $\mathcal{E}_j$ à l'instant k précédent.

La somme des coefficients d'une colonne de A vaut donc 1.

Un telle matrice A que l'on appelle **la matrice de transition** *de la chaine de Markov est dite* **stochastique en colonne.**

Si $A = (a_{i,j})$ et $B = (b_{i,j})$ sont deux matrices stochastiques en colonne d'ordre n, remarquons alors qu'il en est de même de $A \times B$.

En posant en effet $A \times B = (c_{i,j})$, il vient $c_{i,j} = \sum_{k=1}^{n} a_{i,k}.b_{k,j}$.

On a clairement $c_{i,j} \geq 0$, $1 \leq i, j \leq n$.

Puis, pour $j \in [\![1,n]\!]$: $\sum_{i=1}^{n} c_{i,j} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{i,k}.b_{k,j}$

Soit $\sum_{i=1}^{n} c_{i,j} = \sum_{k=1}^{n} b_{k,j}.(\sum_{i=1}^{n} a_{i,k}) = \sum_{k=1}^{n} b_{k,j} = 1$.

Une récurrence sur k montre alors que les puissances A^k, $k \in \mathbb{N}$, de A sont elles-mêmes stochastiques en colonne.

- *Mesurons l'état du système à l'instant k par la variable aléatoire X_k définie de la façon suivante :*

$X_k = i$ si le système est dans l'état $\mathcal{E}_i$ à l'instant $t = k$, $k \geq 0$.

On pose $U_k = \begin{pmatrix} P(X_k = 1) \\ P(X_k = 2) \\ \vdots \\ P(X_k = n) \end{pmatrix}$.

Soit alors $i \in [\![1,n]\!]$.

La formule des probabilités totales appliquée avec le système complet d'événements $(X_k = j)_{1 \leq j \leq n}$ donne :

$$P(X_{k+1} = i) = \sum_{j=1}^{n} P(X_{k+1} = i \mid X_k = j).P(X_k = j) = \sum_{j=1}^{n} p_{i,j}.P(X_k = j).$$

Les égalités précédentes se traduisent par l'égalité matricielle :

$U_{k+1} = A \times U_k$, $k \geq 0$.

De là, on déduit que : $\forall k \in \mathbb{N} : U_k = A^k.U_0$.

Ceci nous amène à considérer la matrice $A^k = (p_{i,j}^{(k)})$.

Supposons que le système soit dans l'état $\mathcal{E}_j$ à l'instant $t = 0$.

La loi de X_0 dans ce cas définie par : $P(X_0 = j) = 1$; $P(X_0 = l) = 0$ si $l \neq j$.

La relation matricielle $U_k = A^k.U_0$ donne :

$$P(X_k = i) = \sum_{l=1}^{n} p_{i,l}^{(k)}.P(X_0 = l) = a_{i,j}^{(k)}.P(X_0 = j) = p_{i,j}^{(k)}.$$

Le coefficient $p_{i,j}^{(k)}$ de la matrice A^k représente donc la probabilité que le système, supposé dans l'état $\mathcal{E}_j$ à l'instant $t = 0$, se trouve dans l'état $\mathcal{E}_i$ à l'instant $t = k$.

On a, de même, la relation matricielle $U_{k+l} = A^l.U_k$.

Supposons la loi de X_l définie par : $P(X_l = j) = 1$; $P(X_l = u) = 0$ si $u \neq j$.

Un raisonnement analogue prouve qu'alors : $P(X_{k+l} = i) = p_{i,j}^{(l)}$.

Le coefficient $p_{i,j}^{(l)}$ représente également la probabilité que le système se trouvant dans l'état $\mathcal{E}_j$ à $t = k$, se trouve ensuite dans l'état $\mathcal{E}_i$ à l'instant $t = k + l$.

- *Supposons maintenant que la suite de matrices $(A^k)_{k \geq 0}$ converge vers une matrice $A^\infty = (p_{i,j}^\infty)$, ce qui signifie que chacune des suites de réels $(p_{i,j}^{(k)})$ converge vers $p_{i,j}^\infty$, $1 \leq i, j \leq n$, lorsque k tend vers $+\infty$.*

On observe la convergence de la suite $(U_k)_{k \geq 0}$ vers une matrice unicolonne U_∞.

De la relation $U_k = A^k.U_0$, on déduit en effet :

$$\forall \in [\![1,n]\!],\ P(X_k = i) = \sum_{l=1}^{n} p_{i,l}^{(k)}.P(X_0 = l) \underset{k \to +\infty}{\longrightarrow} \sum_{l=1}^{n} p_{i,l}^{\infty}.P(X_0 = l).$$

La suite de matrices unicolonne $(U_k)_{k \geq 0}$ converge vers la matrice U_∞ définie par :

$$U_\infty = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \text{ avec } c_i = \sum_{l=1}^{n} p_{i,j}^{\infty}.P(X_0 = l)$$

On reconnait la matrice $A^\infty.U_0$.

On conclut donc que $(U_k)_{k \geq 0} \underset{k \to +\infty}{\longrightarrow} U_\infty = A^\infty.U_0$.

Traduisons maintenant la relation matricielle $A^{k+1} = A.A^k$, $k \geq 0$.

$$\forall (i,j) \in [\![1,n]\!]^2 : p_{i,j}^{(k+1)} = \sum_{l=1}^{n} p_{i,l}.p_{l,j}^{(k)}.$$

Passons à la limite sur k. Il vient : $\forall (i,j) \in [\![1,n]\!]^2,\ p_{i,j}^{\infty} = \sum_{l=1}^{n} p_{i,l}.p_{l,j}^{\infty}$.

On obtient la relation $A^{\infty} = A.A^{\infty}$.

En multipliant les deux membres de cette relation à droite par U_0, il vient :

$A^{\infty}.U_0 = A.(A^{\infty}).U_0 = A.(A^{\infty}.U_0)$.

$U_{\infty} = A.U_{\infty}$.

Les coefficients $c_1, c_2, \ldots, c_n$ de U_{∞} vérifient les propriétés suivantes :

(i) $c_i = \lim_{k \to +\infty} P(X_k = i) \geq 0$.

(ii) $\sum_{i=1}^{n} c_i = \lim_{k \to +\infty} \sum_{i=1}^{n} P(X_k = i) = 1$.

On conclut que ces coefficients $(c_i)_{1 \leq i \leq n}$ sont ceux d'une loi de probabilité.

Soit X_{∞}, une v.a.r. à valeurs dans $[\![1,n]\!]$, admettant cette loi.

$\forall i \in [\![1,n]\!] : P(X_{\infty} = i) = \lim_{k \to +\infty} P(X_k = i)$.

La suite de v.a.r. $(X_k)_{k \geq 0}$ converge en loi vers X_{∞}.

Interprétons maintenant l'égalité $U_{\infty} = A.U_{\infty}$.

Les coefficients de la loi de X_{∞} forment un vecteur propre de A associé à la valeur propre 1.

Au cours du temps, la distribution des probabilités qu'a le système de se trouver dans les différents états tend vers une distribution asymptotique U_{∞} invariante lorsqu'on la multiplie à gauche par la matrice de transition A de la chaine de Markov.

Si cette distribution était atteinte à un instant donné $t = k$, le système la conserverait par la suite.

Dans le cas où cette distribution invariante est celle d'une v.a.r. certaine (les coefficients c_i de U_{∞} étant alors nuls sauf l'un d'eux qui vaut 1), le système est dans l'un des états $\mathcal{E}_i$, $1 \leq i \leq n$ avec la probabilité 1. Il y restera par la suite. Un tel état est appelé **état absorbant**.

- *On voit que l'on est amené naturellement à étudier les matrices stochastiques pour lesquelles une telle distribution invariante U_{∞} existe.*

Pour cela, on cherche des conditions sur A qui impliquent l'existence de la limite matricielle A^{∞}.

On démontre que si tous le coefficents $p_{i,j}$ de A sont strictement positifs, c'est bien le cas.

Le lecteur intéressé pourra se reporter au tome 2- Probabilités et Statistiques de la collection Major - P.U.F.- pour les détails de l'étude (pages 420, 421 et 422).

On y montre que, sous l'hypothèse : $\forall\ (i,j) \in [\![1,n]\!],\ p_{i,j} > 0$, la matrice limite est de la forme : $A^{\infty} = (p_{i,j}^{\infty})$, avec $p_{i,j}^{\infty} = l_i$, $1 \leq i \leq n$, $1 \leq j \leq n$.

Les colonnes de la matrice A^{∞} sont toutes égales à $\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$.

Les réels l_i précédents vérifient $l_i \geq 0$, $1 \leq i \leq n$, et $\sum_{i=1}^{n} l_i = 1$.

De la relation matricielle $U_{\infty} = A^{\infty} \times U_0$, on déduit :

$$P(X_{\infty} = i) = \sum_{j=1}^{n} p_{i,j}^{\infty}.P(X_0 = j) = \sum_{j=1}^{n} l_i.P(X_0 = i) = l_i.\sum_{j=1}^{n} P(X_0 = j) = l_i.$$

La loi asymtotique de X_{∞}, définie par la matrice colonne $U_{\infty} = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$ est indépendante de la distribution au temps $t = 0$ des probabilités entre les différents états, distribution définie par U_0.

Plus généralement, la recherche d'une distribution U de probabilités invariante, c'est-à-dire vérifiant $U = A \times U$ nous conduit à nous intéresser à la valeur propre 1 de A.

Observons que 1 sera valeur propre de A si et seulement si 1 est valeur propre de ${}^tA$.

En effet, les deux matrices $A - \lambda I_n$ et ${}^tA - \lambda.I_n = {}^t(A - \lambda.I_n)$ sont inversibles ou non inversibles en même temps.

Montrons que 1 est valeur propre de ${}^tA$.

Pour cela, $A = (p_{i,j})$ et ${}^tA = (p'_{i,j})$.

Le vecteur $X = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$.

$${}^tA.X = \begin{pmatrix} p'_{1,1} & p'_{1,2} & \cdots & p'_{1,n} \\ p'_{2,1} & p'_{2,2} & \cdots & p'_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ p'_{n,1} & p'_{n,2} & \cdots & p'_{n,n} \end{pmatrix} . \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} p'_{1,j} \\ \sum_{j=1}^{n} p'_{2,j} \\ \vdots \\ \sum_{j=1}^{n} p'_{n,j} \end{pmatrix}.$$

Or : $\sum_{j=1}^{n} p'_{i,j} = \sum_{j=1}^{n} p_{j,i} = 1$

(A est stochastique en colonne).

X est donc vecteur propre de ${}^tA$ associé à la valeur propre $\lambda = 1$.

$\lambda = 1$ est bien valeur propre de A.

Toutefois, le vecteur $X = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ *n'est en général pas vecteur propre de A.*

Le calcul précédent montre qu'il faut, pour cela, que la somme des coefficients de A situés sur une même ligne soit égale à 1.

On dit alors que A est **bistochastique** *(à la fois stochastique en ligne et en colonne).*

- *Qu'en est-il en général de l'existence d'une distribution de probabilités invariante ?*

Si $V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ *est un vecteur propre de A dont toutes les coordonnées sont de même signe, le vecteur*

$U = \dfrac{1}{v_1 + v_2 + \cdots + v_n}.V$ *définit une distribution invariante pour la chaine de Markov.*

Mais l'existence d'un tel vecteur n'est a priori pas assurée. Penchons nous sur cette question.

Pour cela, nous allons procéder par récurrence sur n, et montrer qu'une matrice stochastique A d'ordre n admet toujours un vecteur propre T_n associée à la valeur propre 1, et dont toutes les coordonnées dans la base canonique de $\mathbf{R}^n$ sont ≥ 0.

- *Lorsque $n = 1$, A se réduit au seul coefficient 1, et $T_1 = (1)$ convient.*

- *Supposons la propriété vérifiée pour les matrices stochastiques d'ordre $n-1$, et considérons la matrice stochastique $A = (a_{i,j})$ d'ordre n.*

Ecartons le cas où $A = I_n$, pour lequel le résultat est clair (tout vecteur non nul de $\mathbf{R}^n$ étant vecteur propre de A).

Un vecteur propre $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ associé à la valeur propre 1 de A vérifie le système suivant :

$$\left\{\begin{array}{r} x_1 = a_{1,1}.x_1 + a_{1,2}.x_2 + \cdots + a_{1,n-1}.x_{n-1} + a_{1,n}.x_n \\ \vdots \\ x_{n-1} = a_{n-1,1}.x_1 + a_{n-1,2}.x_2 + \cdots + a_{n-1,n-1}.x_{n-1} + a_{n-1,n}.x_n \\ x_n = a_{n,1}.x_1 + a_{n,2}.x_2 + \cdots + a_{n,n-1}.x_{n-1} + a_{n,n}.x_n \end{array}\right. .$$

Si les coefficients diagonaux $a_{i,i}$ étaient tous égaux à 1, les colonnes de la matrice stochastique A seraient celles de I_n, ce qui est exclu.

Compte tenu de la symétrie des rôles joués par ces coefficients $a_{i,i}$, $1 \leq i \leq n$, on peut supposer $a_{n,n} \neq 1$.

De la dernière équation, on déduit : $x_n = \dfrac{a_{n,1}}{1-a_{n,n}}.x_1 + \cdots \dfrac{a_{n,n-1}}{1-a_{n,n}}.x_{n-1}$.

En reportant dans les $n-1$ premières équations, on aboutit au système suivant :

$$\left\{\begin{array}{r} x_1 = \left(a_{1,1} + \dfrac{a_{1,n}.a_{n,1}}{1-a_{n,n}}\right).x_1 + \cdots + \left(a_{1,n-1} + \dfrac{a_{1,n}.a_{n,n-1}}{1-a_{n,n}}\right).x_{n-1} \\ \vdots \\ x_{n-2} = \left(a_{n-2,1} + \dfrac{a_{n-2,n}.a_{n,1}}{1-a_{n,n}}\right).x_1 + \cdots + \left(a_{n-2,n-1} + \dfrac{a_{n-2,n}.a_{n,n-1}}{1-a_{n,n}}\right).x_{n-1} \\ x_{n-1} = \left(a_{n-1,1} + \dfrac{a_{n-1,n}.a_{n,1}}{1-a_{n,n}}\right).x_1 + \cdots + \left(a_{n-1,n-1} + \dfrac{a_{n-1,n}.a_{n,n-1}}{1-a_{n,n}}\right).x_{n-1} \end{array}\right. .$$

Soit $A' = (a'_{i,j})$, la matrice d'ordre $n-1$ dont le coefficient général est défini par :

$$a'_{i,j} = a_{i,j} + \frac{a_{i,n}.a_{n,j}}{1-a_{n,n}}, \quad 1 \leq i,j \leq n-1.$$

Il vient : (i) $\forall\,(i,j) \in [\![1,n]\!]^2$, $a'_{i,j} \geq 0$.

(ii) $\forall\, j \in [\![1,n]\!]$, $\displaystyle\sum_{i=1}^{n-1} a'_{i,j} = \sum_{i=1}^{n-1}\left(a_{i,j} + \frac{a_{i,n}.a_{n,j}}{1-a_{n,n}}\right) = \sum_{i=1}^{n-1} a_{i,j} + \frac{a_{n,j}}{1-a_{n,n}}.\sum_{i=1}^{n-1} a_{i,n}$.

Soit : $\displaystyle\sum_{i=1}^{n-1} a'_{i,j} = 1 - a_{n,j} + \frac{a_{n,j}}{1-a_{n,n}}.(1-a_{n,n}) = 1$.

La matrice A' est stochastique, d'ordre $n-1$.

Il existe donc un vecteur propre $T_{n-1} = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_{n-1} \end{pmatrix}$ de A' associé à la valeur propre 1 et dont toutes les coordonnées sont ≥ 0.

En posant $x_1 = x'_1$, $x_2 = x'_2$,..., $x_{n-1} = x'_{n-1}$ et $x_n = \dfrac{a_{n,1}}{1-a_{n,n}}.x'_1 + \cdots \dfrac{a_{n,n-1}}{1-a_{n,n}}.x'_{n-1}$, on obtient un vecteur propre T_n de A associé à la valeur propre 1 dont toutes les coordonnées sont ≥ 0.

On a montré par récurrence que toute matrice stochastique possède un vecteur de ce type.

Il suffit ensuite de poser $u_1 = \dfrac{x_1}{\sum_{k=1}^{n} x_k}$, $u_2 = \dfrac{x_2}{\sum_{k=1}^{n} x_k}$,..., $u_n = \dfrac{x_n}{\sum_{k=1}^{n} x_k}$ pour obtenir un vecteur-colonne

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \text{ vérifiant } A \times U = U \text{ et : } \forall\, i \in [\![1,n]\!],\ u_i \geq 0,\ \sum_{i=1}^{n} u_i = 1.$$

Il existe bien une distribution de probabilités invariante pour la chaine de Markov dont A est la matrice de transition.

Etudions maintenant quelques exemples de chaines de Markov. Nous en donnerons à chaque fois la matrice de transition et l'étude asymptotique.

$\boxed{\text{B}}$

$\boxed{\textit{Exemple 1 : Une marche aléatoire}}$

Un mobile se déplace sur trois cases C_1, C_2, C_3 de la façon suivante :

A l'instant $t = 0$, il est en C_1.

Si le mobile est en C_j à un instant $t = k$ quelconque, $k \geq 0$, $1 \leq j \leq 3$, il reste en C_j avec la probabilité $\dfrac{1}{2}$ et passe sur l'une des deux autres cases avec équiprobabilité.

Le système qui décrit la marche aléatoire du mobile peut être dans l'un des trois états suivants :

- Etat $\mathcal{E}_1$: Le mobile est en C_1.

- Etat $\mathcal{E}_2$: Le mobile est en C_2.

- Etat $\mathcal{E}_3$: Le mobile est en C_3.

La matrice de transition A de cette chaine de Markov est donc définie par :

$$p_{i,j} = P(X_{k+1} = i \mid X_k = j) = \frac{1}{4} \text{ si } i \neq j, \text{ et } p_{i,i} = P(X_{k+1} = i \mid X_k = i) = \frac{1}{2},\ 1 \leq i,j \leq 3.$$

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

Pour calculer les puissances de A et passer à la limite, cherchons les éléments propres de A.

- *Appliquons la méthode du pivot de Gauss pour obtenir les valeurs propres de A :*

$$A - \lambda.I_3 = \begin{pmatrix} \frac{1}{2}-\lambda & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2}-\lambda & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{pmatrix} \underset{\substack{L_1 \leftarrow L_1 - 4.(\frac{1}{2}-\lambda).L_3 \\ L_2 \leftarrow L_2 - L_3}}{\longrightarrow} \begin{pmatrix} 0 & \lambda-\frac{1}{4} & (2.\lambda-\frac{1}{2}).(\frac{3}{2}-2.\lambda) \\ 0 & \frac{1}{4}-\lambda & \lambda-\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{pmatrix}$$

$$\underset{\substack{L_1 \leftarrow 1/(\lambda-\frac{1}{4}).L_1 \\ L_2 \leftarrow 1/(\lambda-\frac{1}{4}).L_2}}{\longrightarrow} \begin{pmatrix} 0 & 1 & 3-4.\lambda \\ 0 & -1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{pmatrix}$$

(On a supposé $\lambda \neq \dfrac{1}{4}$ afin d'effectuer la 2ème série d'opérations sur les lignes).

$$\begin{pmatrix} 0 & 1 & 3-4.\lambda \\ 0 & -1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{pmatrix} \underset{L_1 \leftarrow L_1 + L_2}{\longrightarrow} \begin{pmatrix} 0 & 0 & 4-4.\lambda \\ 0 & -1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 4-4.\lambda \\ 0 & -1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{pmatrix} \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \\ 0 & -1 & 1 \\ 0 & 0 & 4.(1-\lambda) \end{pmatrix}.$$

La réduite de Gauss obtenue donne comme seule valeur propre, en dehors du cas $\lambda = \frac{1}{4}$ le réel $\lambda = 1$.

Pour $\lambda = \frac{1}{4}$, la réduite s'écrit :

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac14 & \frac14 & \frac14 \end{pmatrix} \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \begin{pmatrix} \frac14 & \frac14 & \frac14 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Cette réduite n'est pas inversible, et $\lambda = \frac{1}{4}$ est valeur propre de A.

Finalement, les valeurs propres de A sont 1 et $\frac{1}{4}$.

- Cherchons les sous-espaces vectoriels propres de A.

1er cas : $\lambda = 1$.

Le système $A.\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ équivaut au système réduit suivant :

$$\begin{pmatrix} \frac14 & \frac14 & -\frac12 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} . \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

D'où : $\begin{cases} \frac{x}{4} + \frac{y}{4} - \frac{z}{2} = 0 \\ -y + z = 0 \end{cases}$

La solution générale de ce système est (x, y, z), avec $x = y = z$.

En particulier, le vecteur $X_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est un vecteur propre de A associé à la valeur propre $\lambda = 1$.

2ème cas : $\lambda = \frac{1}{4}$.

Le système $A.\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{4}.\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ équivaut au système réduit suivant :

$\begin{pmatrix} \frac14 & \frac14 & \frac14 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, soit à la seule équation : $x + y + z = 0$.

Les vecteurs $X_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $X_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ forment une base de l'espace vectoriel propre associé à $\lambda = \frac{1}{4}$.

(X_2 est non nul, et X_3 est non colinéaire à X_2, sans quoi sa 3ème coordonnée serait nulle).

On conclut que (X_1, X_2, X_3) est une base de $\mathbb{R}^3$ constituée de vecteurs propres de A.

- Soit P, la matrice de passage de la base canonique (e_1, e_2, e_3) de $\mathbb{R}^3$ à la base (X_1, X_2, X_3).

On obtient : $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$.

Pour écrire P^{-1}, exprimons les vecteurs e_j en fonction des vecteurs X_i, $1 \leq i, j \leq 3$.

$$\begin{cases} X_1 = e_1 + e_2 + e_3 \\ X_2 = e_1 - e_2 \\ X_3 = e_1 - e_3 \end{cases}$$

En ajoutant membre à membre les équations, il vient :

$X_1 + X_2 + X_3 = 3.e_1$. D'où $e_1 = \frac{1}{3}.X_1 + \frac{1}{3}.X_2 + \frac{1}{3}.X_3$.

Puis : $e_2 = e_1 - X_2 = \frac{1}{3}.X_1 - \frac{2}{3}.X_2 + \frac{1}{3}.X_3$,

$e_3 = e_1 - X_3 = \frac{1}{3}.X_1 + \frac{1}{3}.X_2 - \frac{2}{3}.X_3$.

On conclut que : $P^{-1} = \frac{1}{3}.\begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$.

Diagonalisons A : $A = P.D.P^{-1}$, avec $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac14 & 0 \\ 0 & 0 & \frac14 \end{pmatrix}$.

Soit $k \in \mathbb{N}$. $A^k = P.D^k.P^{-1}$.

Soit $A^k = \frac{1}{3}.\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} . \begin{pmatrix} 1^k & 0 & 0 \\ 0 & (\frac14)^k & 0 \\ 0 & 0 & (\frac14)^k \end{pmatrix} . \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$.

On trouve finalement : $A^k = \frac{1}{3}.\begin{pmatrix} 1 + \frac{2}{4^k} & 1 - \frac{1}{4^k} & 1 - \frac{1}{4^k} \\ 1 - \frac{1}{4^k} & 1 + \frac{2}{4^k} & 1 - \frac{1}{4^k} \\ 1 - \frac{1}{4^k} & 1 - \frac{1}{4^k} & 1 + \frac{2}{4^k} \end{pmatrix}$.

- Faisons tendre k vers $+\infty$. Il vient :

$A^k \underset{k \to +\infty}{\longrightarrow} A^\infty = \frac13.\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

- La matrice unicolonne U_∞ est définie par la relation $U_\infty = A^\infty.U_0$.

$$U_\infty = \frac{1}{3}.\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} . \begin{pmatrix} P(X_0 = 1) \\ P(X_0 = 2) \\ P(X_0 = 3) \end{pmatrix} = \frac{1}{3}.\begin{pmatrix} P(X_0 = 1) + P(X_0 = 2) + P(X_0 = 3) \\ P(X_0 = 1) + P(X_0 = 2) + P(X_0 = 3) \\ P(X_0 = 1) + P(X_0 = 2) + P(X_0 = 3) \end{pmatrix}.$$

Soit : $U_\infty = \frac{1}{3}.\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

- La loi asymptotique que nous avons étudiée dans le cas général, est ici définie par :

$P(X_\infty = 1) = P(X_\infty = 2) = P(X_\infty = 3) = \frac{1}{3}$.

Il s'agit de la loi uniforme sur $\{1, 2, 3\}$.

Le système tend vers une distribution limite où les trois états sont atteints avec équiprobabilité.

Cela signifie que lorsque le temps s'écoule, le mobile tend à se trouver sur l'une des trois cases de façon équiprobable.

Notons que l'existence de la matrice limite A^∞ était assurée, dans cette situation, car la matrice A avait des coefficients tous strictement positifs.

On peut constater que les colonnes de A^∞ sont égales, comme cela a été déjà dit.

Exemple 2 : Un modèle élémentaire de répartition

On dispose de boules indiscernables en nombre illimité, et de trois boites numérotées de 1 à 3.

On place les boules une à une au hasard dans l'une des trois boites.

Soit X_k, la v.a.r. égale au nombre de boites non vides, après que k boules aient été placées, $k \geq 1$.

Considérons l'évolution des boites au fur et à mesure que l'on place les boules.

Ce nombre de boites non vides varie dans l'ensemble $\{1, 2, 3\}$.

A l'instant $t = k$ (après que l'on ait placé k boules), $k \geq 1$, le système ainsi décrit peut prendre l'un des trois états suivants :

- Etat $\mathcal{E}_1$: il y a une boite non vide.

- Etat $\mathcal{E}_2$: il y a deux boites non vides.

- Etat $\mathcal{E}_3$: il y a trois boites non vides.

- *Relions maintenant les v.a.r. X_{k+1} et X_k, $k \geq 1$.*

La formule des probabilités totales appliquée avec le système complet d'événements $(X_k = j)_{1 \leq j \leq 3}$ donne :

$$P(X_{k+1} = i) = \sum_{j=1}^{3} P(X_{k+1} = i \mid X_k = j).P(X_k = j).$$

(i) Supposons $i = 1$.

$P(X_{k+1} = 1) = P(X_{k+1} = 1 \mid X_k = 1).P(X_k = 1)$

$+P(X_{k+1} = 1 \mid X_k = 2).P(X_k = 2) + P(X_{k+1} = 1 \mid X_k = 3).P(X_k = 3).$

Soit $P(X_{k+1} = 1) = \frac{1}{3}.P(X_k = 1) + 0.P(X_k = 2) + 0.P(X_k = 3).$

Expliquons cette égalité :

$P(X_{k+1} = 1 \mid X_k = 1) = \frac{1}{3}$, *car s'il y a 1 boite non vide après que l'on ait placé k boules, il faut placer la $k+1$ème boule dans cette boite non vide pour qu'il y ait encore une boite non vide ensuite.*

De même : $P(X_{k+1} = 1 \mid X_k = 2) = P(X_{k+1} = 1 \mid X_k = 3) = 0$, *car on ne peut avoir 1 boite non vide avec une $k+1$ème boule si déjà 2 ou 3 des boites sont non vides après que l'on ait placé k boules.*

(ii) Supposons maintenant que $i = 2$.

$P(X_{k+1} = 2) = P(X_{k+1} = 2 \mid X_k = 1).P(X_k = 1)$

$+P(X_{k+1} = 2 \mid X_k = 2).P(X_k = 2) + P(X_{k+1} = 2 \mid X_k = 3).P(X_k = 3).$

Soit $P(X_{k+1} = 2) = \frac{2}{3}.P(X_k = 1) + \frac{2}{3}.P(X_k = 2) + 0.P(X_k = 3).$

Expliquons à nouveau cette égalité.

$P(X_{k+1} = 2 \mid X_k = 1) = \frac{2}{3}$, *car s'il y a 1 boite non vide après que l'on ait placé k boules, il faut placer la $k+1$ème boule dans l'une des 2 boites restées vides pour qu'il y ait 2 boites non vides ensuite.*

$P(X_{k+1} = 2 \mid X_k = 2) = \frac{2}{3}$, *car il faut alors placer la $k+1$ème boule dans l'une des 2 boites non vides pour qu'il y ait encore 2 boites non vides ensuite.*

Enfin : $P(X_{k+1} = 2 \mid X_k = 3) = 0$, *car on ne peut avoir 2 boites non vides avec une $k+1$ème boule si déjà 3 des boites sont non vides après que l'on ait placé k boules.*

Nous laissons au lecteur le soin de justifier la 3ème égalité :

$P(X_{k+1} = 3) = P(X_{k+1} = 3 \mid X_k = 1).P(X_k = 1)$

$+P(X_{k+1} = 3 \mid X_k = 2).P(X_k = 2) + P(X_{k+1} = 3 \mid X_k = 3).P(X_k = 3).$

Soit $P(X_{k+1} = 3) = 0.P(X_k = 1) + \frac{1}{3}.P(X_k = 2) + 1.P(X_k = 3).$

Posons $U_k = \begin{pmatrix} P(X_k = 1) \\ P(X_k = 2) \\ P(X_k = 3) \end{pmatrix}$, $k \geq 1$.

Les trois équations précédemment obtenues s'écrivent :

$$U_{k+1} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix} .U_k = A.U_k.$$

La matrice de transition de la chaine de Markov décrite est donc :

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

- *Il s'agit bien d'une matrice stochastique, mais certains de ses coefficients sont nuls.*

L'existence d'une limite matricielle A^∞ n'est donc pas assurée.

Pour déterminer A^k, $k \in \mathbf{N}$, et passer à la limite, cherchons les éléments propres de A.

- *λ est valeur propre de A si et seulement si la matrice*

$$A - \lambda I_3 = \begin{pmatrix} \frac{1}{3} - \lambda & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} - \lambda & 0 \\ 0 & \frac{1}{3} & 1 - \lambda \end{pmatrix} \text{ est non inversible.}$$

Comme $A - \lambda.I_3$ est triangulaire inférieure, elle est non inversible si et seulement si l'un au moins de ses coefficients diagonaux est nul.

Cela est vérifié si et seulement si $\lambda = \frac{1}{3}$, *ou* $\lambda = \frac{2}{3}$, *ou encore* $\lambda = 1$.

L'ensemble des valeurs propres de A (appelé spectre de A, et noté $Sp(A)$) est donc : $Sp(A) = \{\frac{1}{3}, \frac{2}{3}, 1\}$.

Remarque : *A est une matrice carrée d'ordre 3 et admet 3 valeurs propres. On peut dores et déjà affirmer que A est diagonalisable.*

- *Soient* $X_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $X_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $X_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

$$A.X_1 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix} . \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix} = \frac{1}{3}. \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \frac{1}{3}.X_1.$$

$$A.X_2 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix} . \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} = \frac{2}{3}. \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \frac{2}{3}.X_2.$$

$$A.X_3 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix} . \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = X_3.$$

X_1, X_2, X_3 sont donc des vecteurs propres de A associés respectivement aux valeurs propres $\frac{1}{3}$, $\frac{2}{3}$ et 1.

Comme ils sont associés à des valeurs propres deux à deux distinctes, ils forment une base de $\mathbf{R}^3$.

On retrouve le fait que A est diagonalisable.

- *Soit P, la matrice de passage de la base canonique (e_1, e_2, e_3) de $\mathbf{R}^3$ à la base (X_1, X_2, X_3).*

Il vient : $A = P.D.P^{-1}$, *avec* $D = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *et* $P = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

Déterminons P^{-1}.

Des relations : $\begin{cases} X_1 = e_1 - 2.e_2 + e_3 \\ X_2 = e_2 - e_3 \\ X_3 = e_3 \end{cases}$ *on déduit que :*

$$\begin{cases} e_1 = X_1 + 2.X_2 + X_3 \\ e_2 = X_2 + X_3 \\ e_3 = X_3 \end{cases}.$$

On conclut que : $P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

$$A^k = P.D^k.P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} . \begin{pmatrix} \left(\frac{1}{3}\right)^k & 0 & 0 \\ 0 & \left(\frac{2}{3}\right)^k & 0 \\ 0 & 0 & 1^k \end{pmatrix} . \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Il vient : $A^k = \begin{pmatrix} \frac{1}{3^k} & 0 & 0 \\ 2.\left(\frac{2}{3}\right)^k - \frac{2}{3^k} & \left(\frac{2}{3}\right)^k & 0 \\ 1+\frac{1}{3^k} - 2.\left(\frac{2}{3}\right)^k & 1-\left(\frac{2}{3}\right)^k & 1 \end{pmatrix}$.

- Faisons tendre k vers $+\infty$. On obtient :

$A^k \underset{k\to+\infty}{\longrightarrow} A^\infty = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

Remarque : *Tous les coefficients de A ne sont pas > 0.*

Néanmoins, la limite A^∞ de $(A^k)_{k\geq 0}$ existe.

La condition $\forall\ (i,j) \in [\![1,n]\!]^2,\ a_{i,j} > 0$ est donc une condition suffisante mais non nécessaire pour l'existence de la matrice limite A^∞.

- Ici, $U_\infty = A^\infty.U_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} . \begin{pmatrix} P(X_0=1) \\ P(X_0=2) \\ P(X_0=3) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ P(X_0=1)+P(X_0=2)+P(X_0=3) \end{pmatrix}$.

Soit : $U_\infty = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

La loi asymptotique de cette chaine de Markov est donc définie par :

$P(X_\infty = 1) = P(X_\infty = 2) = 0$ et $P(X_\infty = 3) = 1$.

La suite de v.a.r. $(X_k)_{k\geq 1}$ converge en loi vers la v.a.r. certaine X_∞ égale à 3.

Cela s'explique par le fait que tôt ou tard, les trois boites seront non vides, lorsque l'on y dispose une infinité de boules au hasard.

Exemple 3 : Une urne de Polya

On s'intéresse, dans cette partie, à une suite d'épreuves aléatoires définies de la façon suivante.

Une urne contient des jetons numérotées de 0 à n. Le 1er tirage consiste à choisir au hasard un jeton dans cette urne.

Si le numéro tiré est égal à k, on enlève de l'urne tous les jetons dont le numéro est strictement supérieur à k, et le tirage suivant a lieu dans une urne ne contenant plus que les jetons numérotés de 0 à k.

On poursuit cet algorithme des tirages, et on définit ainsi des v.a.r. X_p par :

$X_p = k$ si le numéro obtenu au pième tirage est k.

- La v.a.r. X_p prend donc ses valeurs dans l'ensemble $\{0,1,2,\ldots,n\}$.

La formule des probabilités totales appliquée avec le système complet d'événements $(X_p = k)_{0\leq k\leq n}$ permet d'écrire :

$$P(X_{p+1} = j) = \sum_{k=0}^{n} P(X_{p+1} = j \mid X_p = k).P(X_p = k).$$

Pour $k < j$, la probabilité conditionnelle $P(X_{p+1} = j \mid X_p = k)$ est nulle.

En effet, si $k < j$, le jeton numéroté j ne figure plus dans l'urne après le pième tirage, et ne peut plus être obtenu par la suite.

Si $k \geq j$, on a : $P(X_{p+1} = j \mid X_p = k) = \dfrac{1}{k+1}$.

En effet, le pième tirage amène alors le numéro k, et j constitue l'un des $k+1$ numéros présents dans l'urne au $(k+1)$ième tirage.

La formule des probabilité totales devient :

$$P(X_{p+1} = j) = \frac{1}{j+1}.P(X_p = j) + \frac{1}{j+2}.P(X_p = j+1) + \cdots + \frac{1}{n+1}.P(X_p = n),\ 0 \leq j \leq n.$$

Posons alors $U_p = \begin{pmatrix} P(X_p = 0) \\ P(X_p = 1) \\ \vdots \\ P(X_p = n) \end{pmatrix}$, $p \geq 1$, et $A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \ldots & \frac{1}{k+1} & \ldots & \frac{1}{n+1} \\ 0 & \frac{1}{2} & \frac{1}{3} & \ldots & \frac{1}{k+1} & \ldots & \frac{1}{n+1} \\ 0 & 0 & \frac{1}{3} & \ldots & \frac{1}{k+1} & \ldots & \frac{1}{n+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 0 & \frac{1}{n+1} \end{pmatrix}$.

L'ensemble des relations précédentes devient équivalent à l'égalité matricielle : $U_{p+1} = A.U_p$, $p \geq 1$.

On remarque que la matrice A est bien stochastique, et l'expérience aléatoire décrite relève encore du formalisme des chaines de Markov.

L'état du système à l'instant $t = p$ est le pième numéro tiré.

Ce système est dans l'état $\mathcal{E}_i$ si le pième numéro tiré est i, $0 \leq i \leq n$.

- La matrice A est triangulaire supérieure. Ses valeurs propres sont ses coefficients diagonaux.

L'ensemble des valeurs propres de A, le spectre de A est donc :

$Sp(A) = \{1, \frac{1}{2}, \ldots, \frac{1}{k+1}, \ldots, \frac{1}{n+1}\}$.

La matrice A, à $n+1$ lignes et $n+1$ colonnes, admet $n+1$ valeurs propres distinctes deux à deux. Elle est donc diagonalisable.

- Numérotons les valeurs propres de A par ordre décroissant, et notons λ_k, la valeur propre $\dfrac{1}{k}$, $1 \leq k \leq n+1$.

Si $(X_1, X_2, \ldots, X_{n+1})$ est une base de vecteurs propres de A, et si P désigne la matrice de passage de la base canonique $(e_1, \ldots, e_{n+1})$ de $\mathbb{R}^{n+1}$ à la base de vecteurs propres précédente, on obtient :

$A = P.D.P^{-1}$, avec $D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_{n+1} \end{pmatrix}$.

D'où : $A^p = P.D^p.P^{-1}$, $p \geq 0$

- Considérons la suite de matrices $(D^p)_{p\geq 0}$.

$D^p = \begin{pmatrix} \lambda_1^p & 0 & \ldots & 0 \\ 0 & \lambda_2^p & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_{n+1}^p \end{pmatrix}$

Comme $0 < \lambda_k < 1$ pour $2 \leq k \leq n+1$, il vient : $(\lambda_k^p) \underset{p\to+\infty}{\longrightarrow} 0$, $(2 \leq k \leq n+1)$.

D'autre part : $\lambda_1 = 1$ et $(\lambda_1^p) \underset{p\to+\infty}{\longrightarrow} 1$.

On conclut que $(D^p) \underset{p\to+\infty}{\longrightarrow} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} = D^\infty$

La matrice limite D^∞ précédente n'a que des coefficients nuls, en dehors du coefficient en position (1,1) qui vaut 1.

- Posons ensuite $P = (\alpha_{i,j})$, $P^{-1} = (\beta_{i,j})$ et $D^p = (d_{i,j}^{(p)})$.

Le coefficient générique $a_{i,j}^{(p)}$ de A^p est alors défini par :

$$a_{i,j}^{(p)} = \sum_{k=1}^{n+1}\sum_{l=1}^{n+1} \alpha_{i,k}.d_{k,l}^{(p)}.\beta_{l,j}.$$

Si $d_{i,j}^{(\infty)}$ désigne la limite de la suite $(d_{i,j}^{(p)})_{p\geq 0}$, on obtient :

$$(a_{i,j}^{(p)}) \underset{p\to+\infty}{\longrightarrow} \sum_{k=1}^{n+1}\sum_{l=1}^{n+1} \alpha_{i,k}.d_{k,l}^{(\infty)}.\beta_{l,j},\ 1 \leq i,j \leq n+1.$$

On conclut que la suite $(A^p)_{p\geq 0}$ admet une limite lorsque p tend vers $+\infty$.

- Le vecteur $X_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ vérifie $A.X_1 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{k+1} & \dots & \frac{1}{n+1} \\ 0 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{k+1} & \dots & \frac{1}{n+1} \\ 0 & 0 & \frac{1}{3} & \dots & \frac{1}{k+1} & \dots & \frac{1}{n+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \frac{1}{n+1} \end{pmatrix} . \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$

Soit : $A.X_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = X_1.$

X_1 est donc un vecteur propre de A associé à la valeur propre 1.

Comme, dans D, la 1ère valeur propre écrite sur la diagonale est 1, on peut choisir X_1 comme premier vecteur de la base de vecteurs propres.

Cherchons la forme de $A^\infty = P.D^\infty.P^{-1}$.

Puisque $X_1 = e_1$, on obtient $e_1 = X_1$, si bien que les 1ères colonnes de P et P^{-1} sont $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

(rappelons que P est la matrice de passage de la base canonique $(e_1, e_2, \dots, e_{n+1})$ de $\mathbf{R}^{n+1}$ à la base de vecteurs propres $(X_1, X_2, \dots X_{n+1})$, et que P^{-1} est la matrice de la base $(X_1, X_2, \dots, X_{n+1})$ à la base $(e_1, e_2, \dots, e_{n+1})$).

Leurs 1ères colonnes sont formées en écrivant respectivement les coordonnées de X_1 dans la base $(e_1, e_2, \dots, e_{n+1})$ et celle de e_1 dans la base $(X_1, X_2, \dots, X_{n+1})$).

Posons alors $P = \begin{pmatrix} 1 & \alpha_{1,2} & \dots & \alpha_{1,n+1} \\ 0 & & & \\ \vdots & & P_1 & \\ 0 & & & \end{pmatrix}$ et $P^{-1} = \begin{pmatrix} 1 & \beta_{1,2} & \dots & \beta_{1,n+1} \\ 0 & & & \\ \vdots & & Q_1 & \\ 0 & & & \end{pmatrix}$, où P_1 et Q_1 désignent des matrices carrées d'ordre n.

Il vient : $D^\infty.P^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} . \begin{pmatrix} 1 & \beta_{1,2} & \dots & \beta_{1,n+1} \\ 0 & & & \\ \vdots & & Q_1 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & \beta_{1,2} & \dots & \beta_{1,n+1} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$

Puis : $P.D^\infty.P^{-1} = \begin{pmatrix} 1 & \alpha_{1,2} & \dots & \alpha_{1,n+1} \\ 0 & & & \\ \vdots & & P_1 & \\ 0 & & & \end{pmatrix} . \begin{pmatrix} 1 & \beta_{1,2} & \dots & \beta_{1,n+1} \\ 0 & 0 & \dots & 0 \\ \dots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$

$P.D^\infty.P^{-1} = \begin{pmatrix} 1 & \beta_{1,2} & \dots & \beta_{1,n+1} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$

A^∞ est bien de la forme : $\begin{pmatrix} b_1 & b_2 & \dots & b_{n+1} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$

avec $b_1 = 1$, $b_2 = \beta_{1,2}, \dots, b_{n+1} = \beta_{1,n+1}$.

- A^∞, limite de la suite de matrices stochastiques A^p, est elle-même stochastique.

La somme des éléments de A^∞, colonne par colonne, vaut donc 1.

Compte-tenu de la forme précédemment trouvée, on obtient :

$b_1 = b_2 = \dots = b_{n+1} = 1$.

On conclut que : $A^{(\infty)} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$

- $U_\infty = A^\infty.U_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} . \begin{pmatrix} P(X_0 = 0) \\ P(X_0 = 1) \\ \vdots \\ P(X_0 = n) \end{pmatrix} = \begin{pmatrix} P(X_0 = 0) + P(X_0 = 1) + \dots + P(X_0 = n) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$

Soit : $U_\infty = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$

La loi asymptotique de la chaine de Markov est ici définie par :

$P(X_\infty = j) = 0$, si $1 \leq j \leq n$, et $P(X_\infty = 0) = 1$.

La v.a.r. X_∞ est donc la v.a.r. certaine égale à 0.

Ceci est naturel, si l'on pense que le numéro 0 sera tiré tôt ou tard, et qu'à partir de cet instant, il n'y aura plus que ce numéro 0 dans l'urne.

- Nous voudrions au passage profiter du formalisme précédent pour montrer au lecteur comment on peut exprimer et déterminer une espérance ou une variance en s'appuyant sur le calcul matriciel.

(i)

- En assimilant une matrice d'ordre 1 à son unique coefficient, on a :

$$E(X_p) = \sum_{k=0}^{n} k.P(X_p = k) = \begin{pmatrix} 0 & 1 & \dots & k & \dots & n \end{pmatrix} . \begin{pmatrix} P(X_0 = 0) \\ P(X_0 = 1) \\ \vdots \\ P(X_0 = n) \end{pmatrix}.$$

Il suffit alors de poser $B = \begin{pmatrix} 0 & 1 & \dots & k & \dots & n \end{pmatrix}$ pour obtenir : $E(X_p) = B.U_p$, $p \geq 1$.

- $B.A$ est une matrice à 1 ligne et $n+1$ colonnes.

Cherchons son $(k+1)$ième coefficient m_{k+1}, $0 \leq k \leq n$.

Pour cela, il suffit de multiplier la ligne de B par la $(k+1)$ième colonne de A.

Cette dernière comprend $k+1$ coefficients égaux à $\frac{1}{k+1}$, suivis de $n-k$ coefficients nuls : $\begin{pmatrix} \frac{1}{k+1} \\ \cdots \\ \frac{1}{k+1} \\ 0 \\ \cdots \\ 0 \end{pmatrix}$.

D'où le $(k+1)$ième coefficient de la matrice unligne $B.A$:

$m_{k+1} = 0.\frac{1}{k+1} + \cdots + k.\frac{1}{k+1} + (k+1).0 + \cdots + n.0 = \frac{1}{k+1}.(1+2+\cdots+k)$.

Soit : $m_{k+1} = \frac{1}{k+1}.\frac{k.(k+1)}{2} = \frac{k}{2}, 0 \leq k \leq n$.

C'est le $(k+1)$ième coefficient de B divisé par 2.

On obtient : $B.A = \frac{1}{2}.B$.

- *La relation $U_{p+1} = A.U_p$ permet d'écrire :*

$B.U_{p+1} = B.(A.U_p) = (B.A).U_p = \frac{1}{2}.B.U_p$.

Soit $E(X_{p+1}) = \frac{1}{2}.E(X_p), p \geq 1$.

- *La suite $(E(X_p)_{p\geq 1}$ est donc une suite géométrique de raison $\frac{1}{2}$.*

D'où : $E(X_p) = (\frac{1}{2})^{p-1}.E(X_1)$.

Lors du 1er tirage, les $n+1$ numéros de 0 à n sont présents, et chacun a la probabilité $\frac{1}{n+1}$ d'être tiré. X_1 suit donc une loi uniforme sur $\{0,\ldots,n\}$.

D'où : $E(X_1) = \frac{0+n}{2} = \frac{n}{2}$.

Finalement : $E(X_p) = (\frac{1}{2})^{p-1}.\frac{n}{2} = \frac{n}{2^p}$.

(ii)

- *Considérons la matrice unligne :* $C = (\, 0^2 \quad 1^2 \quad \ldots \quad k^2 \quad \ldots \quad n^2 \,)$.

On obtient immédiatement :

$$C.U_p = (\, 0^2 \quad 1^2 \quad \ldots \quad k^2 \quad \ldots \quad n^2 \,).\begin{pmatrix} P(X_0 = 0) \\ P(X_0 = 1) \\ \vdots \\ P(X_0 = n) \end{pmatrix}$$

$$C.U_p = \sum_{k=0}^{n} k^2.P(X_p = k) = E(X_p^2), p \geq 1.$$

- *Déterminons le $(k+1)$ième coefficient m'_{k+1} de $C.A$, $0 \leq k \leq n$.*

On l'obtient en multipliant la ligne de C par la $(k+1)$ième colonne de A.

$$\text{Il vient : } m'_{k+1} = (\, 0^2 \quad 1^2 \quad \ldots \quad k^2 \quad \ldots \quad n^2 \,).\begin{pmatrix} \frac{1}{k+1} \\ \vdots \\ \frac{1}{k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \frac{1}{k+1}.(1^2 + 2^2 + \cdots + k^2).$$

Soit : $m'_{k+1} = \frac{1}{k+1}.\frac{k.(k+1).(2.k+1)}{6} = \frac{k^2}{3} + \frac{k}{6}$.

On reconnait le $k+1$ème coefficient de la matrice unligne $\frac{1}{3}.C + \frac{1}{6}.B$.

D'où : $C.A = \frac{1}{3}.C + \frac{1}{6}.B$.

- *Posons alors* $u_p = E(X_p^2) - \frac{n}{2^p}$.

$u_{p+1} = E(X_{p+1}^2) - \frac{n}{2^{p+1}} = C.U_{p+1} + \frac{n}{2^{p+1}}$.

Soit : $u_{p+1} = C.(A.U_p) - \frac{n}{2^{p+1}} = (\frac{1}{3}.C + \frac{1}{6}.B).U_p - \frac{n}{2^{p+1}} = \frac{1}{3}.C.U_p + \frac{1}{6}.B.U_p - \frac{n}{2^{p+1}}$.

Ou encore : $u_{p+1} = \frac{1}{3}.E(X_p^2) + \frac{1}{6}.E(X_p) - \frac{1}{2}.\frac{n}{2^p}$.

Rappelons que $E(X_p) = \frac{n}{2^p}$. Il vient :

$u_{p+1} = \frac{1}{3}.E(X_p^2) - \frac{1}{3}.\frac{n}{2^p} = \frac{1}{3}.u_p$.

- *La suite $(u_p)_{p\geq 1}$ précédente est une suite géométrique de raison $\frac{1}{3}$.*

D'où : $u_p = (\frac{1}{3})^{p-1}.u_1$.

Pour obtenir la valeur de u_1, il nous reste à calculer $E(X_1^2)$.

$X_1 + 1$ suit une loi uniforme sur l'intervalle d'entiers $[\![1, n+1]\!]$.

D'où : $V(X_1+1) = \frac{(n+1)^2-1}{12}$, *et* $V(X_1) = V(X_1+1) = \frac{(n+1)^2-1}{12} = \frac{n^2+2.n}{12}$.

- *La formule de Koënig permet d'écrire :*

$E(X_1^2) = V(X_1) + (E(X_1))^2 = \frac{n^2+2.n}{12} + (\frac{n}{2})^2$.

Soit : $E(X_1^2) = \frac{2.n^2+n}{6}$.

$u_1 = E(X_1^2) - \frac{n}{2} = \frac{n.(n-1)}{3}$.

$u_p = \frac{n.(n-1)}{3^p}, p \geq 1$.

Finalement : $V(X_p) = E(X_p^2) - (E(X_p))^2 = u_p + \frac{n}{2^p} - (\frac{n}{2^p})^2$.

$V(X_p) = \frac{n.(n-1)}{3^p} + \frac{n}{2^p} - \frac{n^2}{4^p}$.

$\boxed{C}$

Nous allons maintenant nous demander succintement s'il est possible de "remonter le temps d'une chaine de Markov" de sorte que l'évolution du système se fasse également selon les modalités d'une chaine de Markov.

- *Pour cela, supposons que toutes les probabilités $P(X_k = i)$, $k \in \mathbb{N}$, $1 \leq i \leq n$ soient strictement positives, et considérons l'égalité :* $P(X_k = i \mid X_{k+l} = j) = \frac{P(X_k = i, X_{k+l} = j)}{P(X_{k+l} = j)}$.

$$P(X_k = i \mid X_{k+l} = j) = \frac{P(X_{k+l} = j, X_k = i)}{P(X_k = i)}.\frac{P(X_k = i)}{P(X_{k+l} = j)} = P(X_{k+l} = j \mid X_k = i).\frac{P(X_k = i)}{P(X_{k+l} = j)}.$$

*Avec les notations de la partie **A**, il vient :*

$$P(X_k = i \mid X_{k+l} = j) = p_{j,i}^{(l)}.\frac{P(X_k = i)}{P(X_{k+l} = j)}.$$

Pour pouvoir parler de chaine de Markov "inverse" au sens donné au **A**, *il faudrait que cette expression ne dépende que de l, et non de k.*

On doit donc pouvoir trouver un n^2-uplet $(u_{1,1,l}, u_{1,2,l}, \cdots, u_{n,n,l})$ de réels tels que :

$\forall k \in \mathbb{N},\ \forall (i,j) \in [\![1,n]\!],\ \dfrac{P(X_k = i)}{P(X_{k+l} = j)} = u_{i,j,l}.$

Pour $l = 1$, il vient : $P(X_{k+1} = j) = \dfrac{P(X_k = i)}{u_{i,j,1}}$.

En prenant $i = j$, on trouve : $P(X_{k+1} = i) = \dfrac{P(X_k = i)}{u_{i,i,1}} = a_i.P(X_k = i)$, *avec* $a_i = \dfrac{1}{u_{i,i,1}}$.

La suite $\left(P(X_k = i)\right)_{k \geq 0}$ est une suite géométrique de raison a_i, et $P(X_k = i) = P(X_0 = i).(a_i)^k$.

La condition $P(X_k = i) \leq 1$ impose $a_i \leq 1$, sans quoi $P(X_0 = i).(a_i)^k$ tendrait vers $+\infty$ avec k.

Sommons ensuite pour i variant de 1 à n. On trouve :

$\displaystyle\sum_{i=1}^{n} P(X_k = i) = 1 = \sum_{i=1}^{n} P(X_0 = i).(a_i)^k.$

Comme les réels a_i^k sont tous ≤ 1, la seule possibilité est $(a_i)^k = 1$, puis $a_i = 1,\ 1 \leq i \leq n$.

On a alors : $\forall\, i \in [\![1,n]\!],\ P(X_k = i) = P(X_0 = i)$.

Les variables X_k doivent avoir toutes même loi.

Posons $u_i = P(X_k = i),\ 1 \leq i \leq n,\ k \in \mathbb{N}$.

La formule de transition de la chaine "directe" : $U_{k+1} = A \times U_k$, *où* $U_{k+1} = \begin{pmatrix} P(X_{k+1} = 1) \\ P(X_{k+1} = 2) \\ \vdots \\ P(X_{k=1} = n) \end{pmatrix}$ *et*

$U_k = \begin{pmatrix} P(X_k = 1) \\ P(X_k = 2) \\ \vdots \\ P(X_k = n) \end{pmatrix}$ *donne :* $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = A.\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$

Le vecteur $U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ *est un vecteur propre de la matrice de transition A associé à la valeur propre 1, et qui a toutes ses coordonnées strictement positives.*

La relation initialement envisagée s'écrit : $P(X_k = i \mid X_{k+l} = j) = p_{j,i}^{(l)}.\dfrac{u_i}{u_j}$.

- *Posons alors* $q_{i,j} = p_{j,i}^{(1)}.\dfrac{u_i}{u_j} = p_{j,i}.\dfrac{u_i}{u_j}$.

Vérifions que la nouvelle matrice $B = (q_{i,j})$ ainsi obtenue est stochastique.

(i) Il est clair que : $\forall\,(i,j) \in [\![1,n]\!]^2,\ q_{i,j} \geq 0$.

(ii) $\displaystyle\sum_{i=1}^{n} q_{i,j} = \sum_{i=1}^{n} p_{j,i}.\frac{u_i}{u_j} = \frac{1}{u_j}.\sum_{i=1}^{n} p_{j,i}.u_i.$

L'égalité matricielle $U = A.U$ implique $\displaystyle\sum_{i=1}^{n} p_{j,i}.u_i = u_j$.

D'où : $\displaystyle\sum_{i=1}^{n} q_{i,j} = \frac{1}{u_j}.u_j = 1.$

- *On peut donc, sous les hypothèses précédentes ($P(X_k = i) > 0$ pour tout $i \in [\![1,n]\!]$ et tout $k \in \mathbb{N}$), définir une chaine de Markov inversée.*

Il faudra que les lois des variables aléatoires X_k soient invariantes par A.

- *Ce faisant, on notera que la transition de l'étape à $t = k+l$ vers l'étape antérieure à $t = k$ s'effectue à l'aide de la matrice B^l (le calcul est analogue à celui mené avec la chaine de Markov directe).*

On a donc les relations : $q_{i,j}^{(l)} = p_{j,i}^{(l)}.\dfrac{u_i}{u_j},\ 1 \leq i,j \leq n,\ l \in \mathbb{N}$.

- *Les transitions décrites par les matrices A et B seront identiques si : $p_{i,j} = q_{i,j}$ quels que soient i et j.*

Cela donne les conditions $p_{i,j} = p_{j,i}.\dfrac{u_i}{u_j},\ 1 \leq i,j \leq n$.

En particulier, si la matrice A est bistochastique, c'est-à-dire à la fois stochastique en ligne et en colonne, le vecteur $X = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ *est un vecteur propre de A associé à la valeur propre 1, et l'on peut choisir* $u_1 = u_2 = \cdots = u_n = \dfrac{1}{n}$ *comme distribution invariante par A.*

La chaine de Markov et sa chaine inverse seront alors identiques si $p_{i,j} = p_{j,i}.\dfrac{u_i}{u_j} = p_{i,j},\ 1 \leq i,j \leq n$, *c'est-à-dire si A est symétrique.*

Dans ce cas A est diagonalisable sur $\mathbb{R}$. Ses valeurs propres sont en valeur absolue inférieures ou égales à 1, et celles dont la valeur absolue est égale à 1 sont des racines $q$ièmes de l'unité pour un entier q compris entre 1 et n. Ces dernières sont donc égales à ± 1.

(pour cette propriété des matrices stochastiques, on pourra se reporter au 2ème problème d'application-Chapitre V- du tome 2 de Probabilités de la collection Major).

- *La marche aléatoire traitée dans l'exemple 1 précédent permet d'envisager la chaine de Markov inverse, car elle admet comme distribution invariante :* $u_1 = u_2 = u_3 = \dfrac{1}{3}$.

Si les variables X_k suivent cette loi uniforme sur $\{1,2,3\}$, on a bien : $P(X_k = i) > 0,\ 1 \leq i \leq 3,\ k \in \mathbb{N}$.

En revanche, l'exemple de la distribution des jetons entre les trois boites de l'exemple 2 ou l'urne de Polya envisagée dans l'exemple 3 ne le permettent pas, car le sous-espace propre associé à la valeur propre 1 de A est de dimension 1, engendré par le vecteur $X = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *pour l'exemple 2 et* $X = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ *pour l'exemple 3.*

La distribution invariante est $u_1 = u_2 = 0$ et $u_3 = 1$ pour l'exemple 2 et $u_0 = 1,\ u_2 = u_3 = \cdots = u_n = 0$ pour l'exemple 3.

Si les variables X_k suivaient cette loi, les coefficients de leur loi ne seraient pas tous strictement positifs.

D

Nous pouvons ensuite prolonger l'étude des chaines de Markov étudiées au **A** *dans une autre direction.*

Le formalisme précédent décrivait les modifications d'un système évoluant entre un nombre fini d'états, les changements d'états se produisant aux instants $t = k,\ k \in \mathbb{N}$, (ou $k \in \mathbb{Z}$ après notre tentative sommaire de remontée du temps).

Or certains processus aléatoires, comme les arrivées de clients à un guichet, se déroulent sur une échelle de temps continue, un client pouvant arriver à tout instant réel t, $t \geq 0$.

Serait-il envisageable de transposer la modélisation précédente pour l'adapter à ce contexte ?

Considérons les variables aléatoires X_k, $k \in \mathbb{N}$ introduites précédemment.

• *Une des conditions qui caractérisaient les chaines de Markov étudiées était la suivante : la transition de l'état $\mathcal{E}_j$ à l'état $\mathcal{E}_i$ entre les instants $t = k$ et $t = k + 1$ ne dépend pas de k et se fait avec la probabilité $p_{i,j}$ si la matrice de transition de la chaine de Markov est $A = (p_{i,j})$.*

Les variables X_k deviennent maintenant des variables aléatoires X_t, $t \geq 0$, indexées par un réel $t \geq 0$ (penser, par exemple, que X_t désigne le nombre de clients qui arrivent entre les instants 0 et t).

L'équivalent naturel de la condition précédente est : $P(X_{s+t} = i \mid X_s = j) = p_{i,j}(t)$, indépendamment de $s \geq 0$.

• *Une autre hypothèse sur ces chaines de Markov consistait à supposer que seul le passé immédiat (l'instant immédiatement antérieur) comptait dans le conditionnement d'un événement lié à l'instant $t = k$.*

Il n'y a évidemment pas d'équivalent direct lorsque l'échelle de temps devient continue. On peut toutefois supposer que ce qui se passe avant un instant s ne compte pas pour juger de ce qui se passera dans l'intervalle de temps $[s, t+s]$, $t \geq 0$, $s > 0$.

Les événements mesurés par les variables aléatoires X_u, $u \geq 0$ seront donc indépendants lorsqu'ils concernent des intervalles de temps de la forme $[s-r, s]$ et $[s, s+t]$, $0 \leq r \leq s \leq t$.

• *Désignons par X_t le nombre d'occurences du phénomène aléatoire observé dans un intervalle de temps d'amplitude t.*

Si $0 \leq u < s < t$, la formule des probabilités totales appliquée avec le système complet d'événements $(X_s = j)$, $j \geq 0$ et la probabilité conditionnelle $P(. \mid X_u = k)$ donne :

$$P(X_t = i \mid X_u = k) = \sum_{j=0}^{+\infty} P(X_t = i \mid X_s = j, X_u = k).P(X_s = j \mid X_u = k).$$

Soit : $P(X_t = i \mid X_u = k)) = \sum_{j=0}^{+\infty} p_{i,j}(t-s).P(X_s = j \mid X_u = k).$

Soit $p_{i,k}(t-u) = \sum_{j=0}^{+\infty} p_{i,j}(t-s).p_{j,k}(s-u).$

Cette équation est connue sous le nom d'équation de **Chapman-Kolmogorov**.

Elle traduit les deux hypothèses précédentes.

• *Pour illustrer cette notion de chaine de Markov continue (on dit* **processus markovien***), étudions en détail le* **processus dit de Poisson***, qui modélise certains phénomènes de files d'attente.*

On envisage l'arrivée de voitures à un péage.

On note alors X_t, la v.a.r. égale au nombre de voitures arrivant entre les instants u et $u+t$, $u, t \geq 0$.

(ce nombre ne dépend que de l'amplitude de l'intervalle de temps), et l'on suppose que $X_t \hookrightarrow \mathcal{P}(a.t)$, $a > 0$ (loi de Poisson de paramètre $a.t$, de sorte que le nombre moyen de voitures arrivant dans un intervalle de temps d'amplitude t est égal à $a.t$, et proportionnel à t).

On suppose ensuite que les voitures arrivent de façon indépendante, lorsque l'on considère des intervalles de temps du type $[u, u+t]$ et $[u+t, u+t+s]$, $u, t, s \geq 0$.

On pose $P_n(t) = P(X_t = n)$, $t \geq 0$, $n \geq 0$.

Etablissons quelques propriétés de ces fonctions : $t \mapsto P_n(t)$.

• $\forall (t, s) \in \mathbb{R}_+ : P_0(t+s) = P_0(t).P_0(s)$.

Notons en effet $A(0; u, v)$, l'événement défini par : "il arrive 0 voiture entre les instants u et v", $u \leq v$.

Soient $s \geq 0$ et $t \geq 0$. Il est clair que $A(0; 0, s+t) = A(0; 0, s) \cap A(0; s, s+t)$.

D'où : $P\big(A(0; 0, s+t)\big) = P\big(A(0; 0, s) \cap A(0; s, s+t)\big) = P\big(A(0; 0, s)\big).P\big(A(0; s, s+t)\big)$

(par hypothèse, les arrivées de voitures dans les intervalles de temps $[0, s]$ et $[s, s+t]$ sont supposées indépendantes).

Soit : $P(X_{s+t} = 0) = P(X_s = 0).P(X_t = 0)$.

On obtient : $P_0(s+t) = P_0(s).P_0(t)$, $0 \leq s \leq< t$.

- $P_1(t) = P(X_t = 1) = e^{-a.t}.\dfrac{(at)^1}{1!} = a.t.e^{-a.t}$.

D'où : $P_1'(t) = a.e^{-a.t} - a^2.t.e^{-a.t}$. Pour $t = 0$, il vient : $P_1'(0) = a$.

- Soit $n \geq 2$. $P_n(t) = e^{-a.t}.\dfrac{(a.t)^n}{n!}$.

$P_n'(t) = -a.\dfrac{(a.t)^n}{n!}.e^{-a.t} + n.a.\dfrac{(a.t)^{n-1}}{n!}.e^{-a.t}$.

D'où : $P_n'(0) = 0$.

- $a.\big(P_{n-1}(t) - P_n(t)\big) = a.e^{-a.t}.(n.\dfrac{(a.t)^{n-1}}{n!} - \dfrac{(a.t)^n}{n!}) = P_n'(t)$.

- $\dfrac{1-P_0(t)}{P_1(t)} - 1 = \dfrac{1-P_0(t)-P_1(t)}{P_1(t)} = \dfrac{1-e^{-a.t}-a.t.e^{-a.t}}{a.t.e^{-a.t}} = \dfrac{e^{a.t}-1-a.t}{a.t} = \dfrac{o(t)}{t} = o(1)$.

On conclut que $\dfrac{1-P_0(t)-P_1(t)}{P_1(t)} \underset{t\to 0^+}{\longrightarrow} 0$, puis que $\dfrac{1-P_0(t)}{P_1(t)} \underset{t\to 0^+}{\longrightarrow} 1$.

Examinons maintenant une réciproque qui donne des conditions caractéristiques de ces processus de Poisson.

On suppose pour cela que l'arrivée des voitures au péage répond aux hypothèses suivantes :

(H_1) : Le nombre de voitures arrivant dans un intervalle de temps donné ne dépend que de l'amplitude de cet intervalle. Il est nul si l'amplitude est nulle

(H_2) : Les voitures arrivent de façon indépendante dans des intervalles de temps du type $[u, u+t]$ et $[u+t, u+t+s]$, $u \geq 0, t \geq 0, s \geq 0$.

(H_3) : Si $P_n(t)$ désigne la probabilité qu'il arrive n voitures entre les instants 0 et t, alors la fonction $t \mapsto P_n(t)$ est dérivable sur $\mathbf{R}_+$.

(H_4) : $\dfrac{1-P_0(t)}{P_1(t)} \underset{t\to 0^+}{\longrightarrow} 1$.

Posons $f(t) = P_0(t)$, $t \geq 0$ et déterminons la forme de f.

- f est décroissante.

En effet, si $A(0;0,t)$ désigne l'événement : "il n'arrive pas de voiture entre les instants 0 et t".

Il est clair que $A(0;0,t')$ implique $A(0;0,t)$ pour $t \leq t'$. En effet, s'il n'arrive pas de voiture entre les instants 0 et t', il n'a pu en arriver entre les instants 0 et t.

D'où : $P\big(A(0;0,t')\big) \leq P\big(A(0;0,t)\big)$.

Lorsque $0 \leq t \leq t'$, on obtient : $f(t') \leq f(t)$. f est décroissante sur $\mathbf{R}_+$.

- $\forall (s,t) \in \mathbf{R}+^2 : f(s+t) = f(s).f(t)$.

Justifions cette égalité.

$f(s+t) = P\big(A(0;0,s+t)\big) = P\big(A(0;0,s) \cap A(0;s,s+t)\big)$

(il n'arrive aucune voiture entre les instants 0 et $s+t$ s'il n'en arrive aucune entre les instants 0 et s, et aucune entre les instants s et $s+t$).

Les événements $A(0;0,s)$ et $A(0;s,s+t)$ sont indépendants, d'après l'hypothèse (H_2).

D'où : $\forall (s,t) \in \mathbf{R}_+^2 : f(s+t) = P\big(A(0;0,s)\big).P\big(A(0;s,s+t)\big) = f(s).f(t)$.

- Cherchons les valeurs possibles de $f(0)$.

Pour $s = t = 0$, l'égalité précédente donne : $f(0) = f(0)^2$, soit $f(0) = 0$ ou $f(0) = 1$.

$f(0) = 0$ signifierait que $P\big(A(0;0,0)\big) = 0$.

Cela supposerait que l'on a une probabilité nulle de voir arriver 0 voiture entre les instants 0 et 0.

Au contraire $P\big(A(0;0,0)\big) = 1$. D'où : $f(0) = 1$.

- Une récurrence simple sur n montre ensuite que : $\forall x > 0$, $\forall n \in \mathbf{N}$: $f(n.x) = \big(f(x)\big)^n$.

- Supposons que $f(1) = 0$.

Soit $x > 0$. Il existe $n \geq 1$ tel que $n.x > 1$.

La décroissance de f implique $0 \leq f(n.x) \leq f(1) = 0$, soit $f(n.x) = 0$. D'où : $\big(f(x)\big)^n = 0$ et $f(x) = 0$, $x > 0$.

Il viendrait : $f(0) = \lim\limits_{x\to 0^+} f(x) = 0$, alors que $f(0) = 1$ (probabilité qu'il arrive 0 voiture entre les instants 0 et 0).

On conclut donc que $0 < f(1) \leq f(0) = 1$.

- Envisageons maintenant l'hypothèse $f(1) = 1$.

Elle signifierait qu'entre les instants 0 et 1, il n'arrive aucune voiture de façon presque certaine.

On pourrait écrire : $P\big(A(0;0,1)\big) = P\big(A(0;1,2)\big) = \cdots = P\big(A(0;n-1,n)\big) = 1$.

De là, on déduirait :

$$P\big(\overline{A(0;0,n)}\big) = P\big(\overline{\bigcap_{k=1}^{n} A(0;k-1,k)}\big) = P\big(\bigcup_{k=1}^{n} \overline{A(0;k-1,k)}\big) \leq \sum_{k=1}^{n} P\big(\overline{A(0;k-1,k)}\big) = 0.$$

D'où : $P\big(\overline{A(0;0,n)}\big) = 0$, puis : $P\big(A(0;0,n)\big) = 1$, $n \geq 1$.

La suite $\big(A(0;0,n)\big)_{n\geq 1}$ d'événements est décroissante $(A(0;0,n+1) \subset A(0;0,n)$ car s'il n'arrive aucune voiture entre les instants 0 et $n+1$, il n'en arrive a fortiori aucune entre les instants 0 et n).

On conclurait donc que : $P\big(A(0;0,n)\big) \underset{n\to+\infty}{\longrightarrow} P\big(\bigcap_{n=0}^{+\infty} A(0;0,n)\big)$.

D'où l'égalité : $P\big(\bigcap_{n=0}^{+\infty} A(0;0,n)\big) = 1$.

Il n'arriverait jamais aucune voiture de façon presque certaine. Le péage serait bien désert !!

Nous écarterons ce cas, si bien que $f(1) < 1$.

- Il existe alors $a > 0$ tel que $f(1) = e^{-a}$.

Une nouvelle récurrence sur n prouve que : $\forall n \in \mathbf{N}$, $f(n) = e^{-a.n}$.

- Déterminons ensuite f sur les fractions rationnelles.

Posons $r = \dfrac{p}{q}$, $p \geq 1$, $q \geq 1$.

$f(q.r) = \big(f(r)\big)^q$, d'après la règle établie précédemment.

D'autre part : $f(q.r) = f(p) = e^{-a.p}$.

D'où : $f(r) = \sqrt[q]{e^{-a.p}} = e^{-a.\frac{p}{q}} = e^{-a.r}$.

- Déterminons enfin f sur $\mathbf{R}_+$.

Soit x, un réel ≥ 0.

Notons $(r_n)_{n\geq 0}$ et $(t_n)_{n\geq 1}$, les suites des approximations décimales de x par excès et par défaut.

$r_n \leq x < t_n$, $n \geq 0$. *La décroissance de f permet d'écrire :* $f(t_n) \leq f(x) \leq f(r_n)$.

Soit : $e^{-a.t_n} \leq f(x) \leq e^{-a.r_n}$.

Faisons tendre n vers $+\infty$. $\lim\limits_{n\to+\infty} e^{-a.t_n} = \lim\limits_{\to+\infty} e^{-a.r_n} = e^{-a.x}$.

On conclut que $f(x) = e^{-a.x}$, $x \in \mathbb{R}_+^*$.

- $\dfrac{1-P_0(t)}{P_1(t)} = \dfrac{1-e^{-a.t}}{P_1(t)} \underset{+\infty}{\approx} \dfrac{a.t}{P_1(t)}$.

D'où : $\dfrac{P_1(t)-P_1(0)}{t} = \dfrac{P_1(t)}{t} \underset{+\infty}{\approx} a.\dfrac{P_1(t)}{1-P_0(t)} \underset{x\to0^+}{\longrightarrow} a.1 = a$

(revoir l'hypothèse (H_4)).

La fonction P_1 est dérivable à droite en 0, et $P_1{}'(0) = a$.

De même, les événements $A(n;0,t)$, $n \geq 0$, forment un système complet (entre les instants 0 et t, il arrive un certain nombre n de voitures).

$\sum\limits_{n=0}^{+\infty} P(A(n;0,t)) = 1$.

Soit n, un entier ≥ 2. Il vient $P(A(0;0,t)) + P(A(1;0,t)) + P(A(n;0,t)) \leq 1$.

D'où : $0 \leq \dfrac{P_n(t)-P_n(0)}{t} = \dfrac{P_n(t)}{t} \leq \dfrac{1-P_0(t)-P_1(t)}{t}$.

Comme $\dfrac{P_1(t)}{t} \underset{+\infty}{\approx} a$, *il vient* $t \underset{+\infty}{\approx} \dfrac{P_1(t)}{a}$.

D'où : $\dfrac{1-P_0(t)-P_1(t)}{t} \underset{+\infty}{\approx} a.\dfrac{1-P_0(t)-P_1(t)}{P_1(t)} = a.(\dfrac{1-P_0(t)}{P_1(t)} - 1)$,

et $a.(\dfrac{1-P_0(t)}{P_1(t)} - 1) \underset{t\to0^+}{\longrightarrow} 0$.

On conclut que $\dfrac{P_n(t)-P_n(0)}{t} \underset{t\to0^+}{\longrightarrow} 0$.

P_n est dérivable à droite en 0, et $P_n{}'(0) = 0$, $n \geq 2$.

- *Etablissons maintenant l'égalité :* $P_n(t+s) = \sum\limits_{k=0}^{n} P_{n-k}(t).P_k(s)$

Pour cela, notons qu'il arrive n voitures entre les instants 0 et $s+t$, s'il en arrive k, $0 \leq k \leq n$, entre les instants 0 et s, et $n-k$ entre les instants s et $s+t$.

$A(n;0,s+t) = \bigcup\limits_{k=0}^{n} A(k;0,s) \cap A(n-k;s,s+t)$.

$P\big(A(n;0,s+t)\big) = \sum\limits_{k=0}^{n} P\big(A(k;0,s) \cap A(n-k;s,t)\big)$.

Les événements $A(k;0,s)$ et $A(n-k;s,s+t)$ étant indépendants d'après l'hypothèse (H_2), il vient :

$P\big(A(n;0,s+t)\big) = \sum\limits_{k=0}^{n} P\big(A(k;0,s)\big).P\big(A(n-k;s,s+t)\big)$.

Soit : $P_n(s+t) = \sum\limits_{k=0}^{n} P_k(s).P_{n-k}(t)$.

- *De l'égalité* $P_n(s+t) = \sum\limits_{k=0}^{n} P_{n-k}(t).P_k(s)$, *on déduit :*

$P_n(t+s) = P_n(t).P_0(s) + \sum\limits_{k=1}^{n} P_{n-k}(t).P_k(s)$.

Soit : $P_n(t+s) - P_n(t) = P_n(t).(e^{-a.s} - 1) + \sum\limits_{k=1}^{n} P_{n-k}(t).P_k(s)$.

$\dfrac{P_n(t+s)-P_n(t)}{s} = P_n(t).\dfrac{e^{-a.s}-1}{s} + \sum\limits_{k=1}^{n} P_{n-k}(t).\dfrac{P_k(s)-P_k(0)}{s}$

$(P_k(0) = 0$ *pour* $k \geq 1)$.

Faisons tendre s vers 0^+.

$\dfrac{e^{-a.s}-1}{s} \underset{s\to0^+}{\longrightarrow} -a$.

$\dfrac{P_k(s)-P_0(s)}{s} \underset{s\to0^+}{\longrightarrow} P_k'(0) = \begin{vmatrix} a & si & k=1 \\ 0 & si & k>1 \end{vmatrix}$

D'où : $\dfrac{P_n(t+s)-P_n(t)}{s} \underset{s\to0^+}{\longrightarrow} P_n(t).(-a) + P_{n-1}(t).a$.

$\dfrac{P_n(t+s)-P_n(t)}{s} \underset{s\to0^+}{\longrightarrow} a.(P_{n-1}(t) - P_n(t))$.

P_n est dérivable en t (à droite), et $P_n{}'(t) = a.(P_{n-1}(t) - P_n(t))$.

- *Posons maintenant* $Q_n(t) = e^{a.t}.P_n(t)$, $t \geq 0$.

Il vient : $Q_n'(t) = a.e^{a.t}.P_n(t) + e^{a.t}.P_n'(t) = a.e^{a.t}.\big(P_n(t) + P_{n-1}(t) - P_n(t)\big)$.

Soit : $Q_n'(t) = a.e^{a.t}.P_{n-1}(t) = a.Q_{n-1}(t)$.

Raisonnons ensuite par récurrence sur n, pour montrer que : $\forall n \in \mathbb{N} : Q_n(t) = \dfrac{(a.t)^n}{n!}$.

(i) Pour $n = 0 : Q_0(t) = e^{a.t}.P_0(t) = e^{a.t}.e^{-a.t} = 1$.

(ii) Supposons que $Q_{n-1}(t) = \dfrac{(a.t)^{n-1}}{(n-1)!}$, $t \geq 0$, $n \geq 1$.

$Q_n'(t) = a.Q_{n-1}(t) = a.\dfrac{(a.t)^{n-1}}{(n-1)!}$.

D'où : $Q_n(t) = \dfrac{(a.t)^n}{n!} + k_n$, $k_n \in \mathbb{R}$, $(n \geq 1)$.

Pour $t = 0 : Q_n(t) = a.P_n(t) = 0$.

On conclut que : $Q_n(t) = \dfrac{(a.t)^n}{n!}$.

On remarquera que la formule donnant $Q_n(t)$ reste valable pour $n = 0$.

Puis, du calcul précédent, on déduit que : $P_n(t) = e^{-a.t}.\dfrac{(a.t)^n}{n!}$, $(t \in \mathbb{R}_+, n \geq 0)$.

Pour finir, soit X_t, la v.a.r. égale au nombre de voitures arrivant au péage pendant un intervalle de temps d'amplitude t.

$P(X_t = n) = P_n(t) = e^{-a.t}.\dfrac{(a.t)^n}{n!}$, $n \geq 0$.

X_t suit une loi de Poisson de paramètre $a.t$: $X_t \hookrightarrow \mathcal{P}(a.t)$.

$E(X_t) = a.t$ *(et* $V(X_t) = a.t$*).*

Le modèle précédent, avec les hypothèses (H_1) à (H_4) redonne donc le modèle où les varaibles aléatoires X_t sont des variables de Poisson telles que le nombre moyen de voitures arrivant au péage soit proportionnel à l'amplitude t de l'intervalle d'observation.

Un tel processus markovien est naturellement appelé **processus de Poisson**.