

DM 22 - Chaînes de Markov

1 Théorème de Perron-Frobenius

1. Pour tout $i \in \llbracket 1, n \rrbracket$, $(SX)_{i,1} = \sum_{j=1}^n S_{i,j} x_j = \sum_{j=1}^n S_{i,j}$.

On a donc $SX = X$ ssi $\forall i \in \llbracket 1, n \rrbracket$, $\sum_{j=1}^n S_{i,j} = 1$ ssi S est stochastique.

2. Soient S, T stochastiques. Comme les coefficients de S et T sont positifs, ceux de ST aussi. De plus, on a $(ST)X = S(TX) = SX = X$. Donc, par la question précédente, ST est stochastique.

3. • On considère la i -ème composante de l'égalité $SY = \lambda Y$. On a $(SY)_i = \sum_{j=1}^n S_{i,j} y_j$ et $(\lambda Y)_i = \lambda y_i$. On prend la valeur absolue et on a la première égalité.
- On applique ensuite l'inégalité triangulaire en utilisant $S_{i,j} \geq 0$.
- La deuxième inégalité vient de ce qu'on a pris $|y_i|$ maximal et qu'on a ensuite normalisé pour supposer $|y_i| \leq 1$. Donc, on a $|y_j| \leq 1$ pour tout j .
- La dernière égalité vient de la définition d'une matrice stochastique.

4. La chaîne d'(in)égalités montre que $|\lambda| \leq 1$. Si $|\lambda| = 1$, on a égalité dans les deux inégalités. L'égalité dans la deuxième inégalité donne $|y_j| = 1$ pour tout j (car S est supposée strictement positive). L'égalité dans la première inégalité donne qu'il existe $u \in \mathbb{C}$ et des $\mu_j \in \mathbb{R}_+$ tels que $S_{i,j} y_j = \mu_j u$. Comme les $S_{i,j}$ sont des réels positifs, les $\frac{y_j}{u}$ sont des réels positifs. Donc les y_j sont de module 1 et ont même argument. Donc $Y \in \text{Vect}(X)$.

5. S^p est strictement positive par hypothèse. Soient $i, k \in \llbracket 1, n \rrbracket$. On a

$$(S^{p+1})_{i,k} = \sum_{j=1}^n S_{i,j} (S^p)_{j,k}.$$

Tous les $(S^p)_{j,k}$ sont strictement positifs ; de plus les $S_{i,j}$ sont positifs et au moins un est non nul puisque leur somme vaut 1 ; donc S^{p+1} est strictement positive. De plus, par Q2, S^p et S^{p+1} sont stochastiques.

Comme $SX = \lambda X$, on a $S^p X = \lambda^p X$ et $S^{p+1} X = \lambda^{p+1} X$. Donc, λ^p est une valeur propre de S^p et λ^{p+1} de S^{p+1} . On peut donc appliquer les résultats de la question précédente : on a $|\lambda|^p \leq 1$, donc $|\lambda| \leq 1$.

De plus, si $|\lambda| = 1$, on a $|\lambda^p| = |\lambda^{p+1}| = 1$. Par le cas d'égalité de la question précédente, $\lambda^p = \lambda^{p+1} = 1$, donc $\lambda = 1$. De plus, $Y \in \text{Vect}(X)$ (car c'est un vecteur propre associé à λ^p pour la matrice S^p).

6. Comme les matrices λI_n et N commutent, on peut appliquer la formule du binôme de Newton.

$$T^p = \sum_{k=0}^p \binom{p}{k} \lambda^{p-k} N^k.$$

De plus, comme N est nilpotente, les termes sont nuls si $p \geq n$. Fixons $k \in [0, n-1]$. La quantité $\binom{p}{k}$ est polynomiale en p , tandis que λ^{p-k} est une suite géométrique de raison λ avec $|\lambda| < 1$. Par croissance comparée, on a $\binom{p}{k} \lambda^{p-k} \rightarrow 0$ quand $p \rightarrow +\infty$. Il en va de même des matrices $\binom{p}{k} \lambda^{p-k} N^k$. Par somme finie, T^p converge vers la matrice nulle.

7. Les puissances S'^p se calculent par un produit par blocs :

$$S'^p = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & A_1^p & 0 & \dots & 0 \\ 0 & 0 & A_2^p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & A_r^p \end{pmatrix}.$$

Par la question précédente, chaque bloc A_k^p tend vers 0. Donc, S'^p tend vers la matrice K avec un 1 en haut à gauche, et des 0 partout ailleurs.

8. Comme S est semblable à S' , on peut trouver une matrice $Q \in GL_n(\mathbb{R})$ telle que $S = QS'Q^{-1}$. On en déduit que $S^p = QS'^pQ^{-1}$, pour tout p . On montre (en regardant coefficient par coefficient) que les limites matricielles sont compatibles avec le produit matriciel. Donc, S^p converge vers QKQ^{-1} , où K est définie à la question précédente.

9. Pour tout p et pour tous i, j , $(S^p)_{i,j} \geq 0$. Par passage à la limite, on en déduit que $(S_\infty)_{i,j} \geq 0$.

De plus, pour tout i , $\sum_{j=1}^n (S^p)_{i,j} = 1$; donc par passage à la limite $\sum_{j=1}^n (S_\infty)_{i,j} = 1$.

Donc S_∞ est stochastique. Elle est de rang 1 car semblable à la matrice K qui est de rang 1.

10. Comme S_∞ est de rang 1, elle a une ligne non nulle $L = (p_1 \ p_2 \ \dots \ p_n)$ et toutes les lignes sont proportionnelles. De plus, comme S_∞ est stochastique, la somme sur chaque ligne vaut 1 ; donc $\sum_{j=1}^n p_j = 1$ et toutes les autres lignes sont égales à L (car sinon la somme ne ferait pas

1). Ainsi, S_∞ est de la forme

$$S_\infty = \begin{pmatrix} p_1 & p_2 & \dots & p_n \\ p_1 & p_2 & \dots & p_n \\ \vdots & \vdots & \vdots & \vdots \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

ce qu'on peut encore écrire $S_\infty = XL$.

11. On a $S_\infty S = S_\infty$ par passage à la limite dans l'égalité $S^p S = S^{p+1}$. Donc, $XL S = XL$. C'est une égalité entre deux matrices ayant leurs n lignes identiques ; à gauche, la ligne vaut LS , à droite L . Donc, $LS = L$.

Si $\tilde{L} S = \tilde{L}$, on a par récurrence immédiate $\tilde{L} S^p$, puis $\tilde{L} S_\infty = \tilde{L}$ par passage à la limite. On a donc $(\tilde{L} X) L = \tilde{L}$. Comme $\tilde{L} X \in \mathbb{R}$, ceci montre que $\tilde{L} \in \text{Vect}(L)$.

2 Théorème ergodique sur les chaînes de Markov

12. Par la formule des probabilités composées, on a

$$P((X_0 = i_0) \cap \cdots \cap (X_p = i_p)) = P(X_0 = i_0) \times \prod_{k=1}^p P(X_k = i_k \mid (X_0 = i_0) \cap \cdots \cap (X_{k-1} = i_{k-1})).$$

Par hypothèse markovienne et indépendance temporelle, les facteurs se simplifient en $P(X_k = i_k \mid X_{k-1} = i_{k-1}) = P(X_1 = i_k \mid X_0 = i_0) = P(i_{k-1}, i_k)$. De plus, $P(X_0 = i_0) = \pi_0(i_0)$. D'où la formule annoncée.

13. Soit $i \in \llbracket 1, n \rrbracket$. On a

$$\sum_{j=1}^n P(i, j) = \sum_{j=1}^n P(X_1 = j \mid X_0 = i).$$

Les évènements $(X_1 = j)$, pour $j \in \llbracket 1, n \rrbracket$ forment un système complet d'évènements. Donc,

$$\sum_{j=1}^n P(X_1 = j \mid X_0 = i) = 1$$

On rappelle que si A est un évènement quelconque de probabilité non nulle, P_A est une probabilité ; on applique cela à l'évènement $(X_0 = i)$.¹

14. Soit $i \in \llbracket 1, n \rrbracket$. On a

$$P(X_{p+1} = i) = \sum_{j=1}^n P(X_p = j)P(j, i),$$

en utilisant Q12. Cette égalité se traduit matriciellement en $\pi_{p+1} = \pi_p P$.

15. Si ν est invariante, on a $\nu P = \nu$ car $\nu = \pi_0$ et $\nu_p = \pi_1$.

Réciproquement, si $\nu P = \nu$, une récurrence immédiate sur la question précédente donne $\pi_p = \nu P^p = \nu$, de nouveau par récurrence.

16. Pour tout p , la matrice P^p a pour coefficient (i, j) la valeur $P(X_p = j \mid X_0 = i)$. De plus, on sait que P est stochastique. Dès lors, les conditions de forte irréductibilité et d'ergodicité sont les mêmes.

17. On applique les question 10 et 11 sur la matrice P ergodique. En notant ν la ligne appelée L question 10, on a $\nu P = P$. De plus, la somme des coefficients de ν vaut 1, donc ν peut être interprétée comme une mesure de probabilité, et elle est invariante (Q15).

L'unicité vient de ce que si ν' est invariante, on a $\nu' P = P$ et donc $\nu' \in \text{Vect}(\nu)$ (Q11) et donc $\nu' = \nu$ car les sommes des coefficients valent 1.

On a vu que $\pi_p = \pi_0 P^p$. Le terme de droite tend vers $\pi_0 P_\infty = \pi_0 X \nu$, où X est la matrice colonne avec des 1. Mais, $\pi_0 X = 1$ (la somme des coefficients de π_0 vaut 1) ; on a donc montré que π_p converge vers ν .

3 Deux exemples

3.1 Chaîne de Markov à deux états

18. On a $Q = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$.

¹Formellement, rien n'assure que $(X_0 = i)$ est de probabilité non nulle ; il aurait fallu préciser cela dans l'énoncé ou présenter les choses légèrement différemment.

19. Comme $\alpha, \beta \in]0, 1[$, Q est strictement positive donc ergodique. Donc, la chaîne de Markov est fortement irréductible. Pour déterminer la loi invariante $\nu = (\nu_1 \quad \nu_2)$, on résout le système $\nu Q = \nu$. On trouve $\nu \in \text{Vect}((\beta \quad \alpha))$. Donc, après normalisation :

$$\nu_1 = \frac{\beta}{\alpha + \beta} \text{ et } \nu_2 = \frac{\alpha}{\alpha + \beta}.$$

Et Q^p converge vers $\begin{pmatrix} \frac{\beta}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \\ \frac{\beta}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \end{pmatrix}$.

3.2 Urnes d'Ehrenfest

20. On a une chaîne à $n + 1$ états qu'on note par les entiers de $\llbracket 0, n \rrbracket$. On est dans l'état k s'il y a k boules dans l'urne 1. Quand on est dans l'état, on peut seulement aller dans l'état $k + 1$ ou $k - 1$ (si on ne sort pas de $\llbracket 0, n \rrbracket$). La probabilité d'aller de k à $k + 1$ est $\frac{n - k}{n}$ et celle d'aller de k à $k - 1$ est $\frac{k}{n}$.

21. On note P cette matrice (avec les lignes et colonnes de 0 à n). On a

$$P = \frac{1}{n} \begin{pmatrix} 0 & n & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & n-1 & 0 & \dots & 0 & 0 \\ 0 & 2 & 0 & n-2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

22. La parité du nombre de boules dans une urne évolue à chaque étape. Cela se traduit par le fait que les matrices P^p ont des 0 en position (i, j) si p est pair et que i et j n'ont pas même parité ; ou si p est impair et que i et j ont même parité.

Donc, aucune puissance de P n'est strictement positive et la chaîne de Markov n'est pas fortement irréductible.

23. Le vecteur ligne ν correspondant à la loi binomiale a pour coefficient k : $\frac{\binom{n}{k}}{2^n}$. Pour $k \in \llbracket 1, n - 1 \rrbracket$, le coefficient k de νP est

$$\frac{1}{n2^n} \left((n - k + 1) \binom{n}{k-1} + (k + 1) \binom{n}{k+1} \right).$$

Par symétrie et la formule du chef, cela se simplifie en

$$\frac{1}{n2^n} \left(n \binom{n-1}{n-k} + n \binom{n-1}{k} \right) = \frac{1}{2^n} \binom{n}{k},$$

par la formule de Pascal (*Les trois formules ensemble !*) Ceci montre l'égalité $\nu P = \nu$ pour $k \neq 0, n$ (et les cas restants sont immédiats).

Pour montrer l'unicité, on écrit le système linéaire correspondant et on observe assez vite que l'ensemble des solutions est de dimension 1 (*à écrire*).

24. En général, non. Si la loi de X_0 favorise les états pairs par exemple, la loi de X_p va favoriser les états pairs quand p est pair et les états impairs quand p est impair. Ceci implique que la loi de X_p ne va pas converger.