DM 22 - Chaînes de Markov

1 Théorème de Perron-Frobenius

Si A et M sont des matrices de $\mathcal{M}_n(\mathbb{C})$, on dit que la suite $(A^p)_{p\in\mathbb{N}}$ converge vers M si, pour tous indices $i, j \in [\![1, n]\!]$, $(A^p)_{i,j} \to M_{i,j}$, quand $p \to +\infty$.
On note $X \in \mathcal{M}_{n,1}(\mathbb{R})$ la matrice colonne dont les coefficients sont égaux à 1.
Une matrice S de $\mathcal{M}_n(\mathbb{R})$ est

- *positive* si $\forall i, j \in [\![1, n]\!], S_{i,j} \geq 0$;
- *strictement positive* si $\forall i, j \in [\![1, n]\!], S_{i,j} > 0$;
- *stochastique* si elle est positive et si $\forall i \in [\![1, n]\!], \displaystyle\sum_{j=1}^{n} S_{i,j} = 1$;
- *ergodique* si elle est stochastique et s'il existe $p \in \mathbb{N}^*$ tel que S^p est strictement positive.

L'objectif de cette partie est de montrer la convergence de la suite S^n quand S est ergodique et de décrire précisément la matrice limite S_∞. C'est une version du théorème de Perron[1]-Frobenius[2].

1. Soit $S \in \mathcal{M}_n(\mathbb{R})$ une matrice positive. Montrer que S est stochastique ssi $SX = X$.
2. En déduire que l'ensemble des matrices stochastiques est stable par produit matriciel.

On suppose que S est stochastique et strictement positive ; on considère S comme une matrice dans $\mathcal{M}_n(\mathbb{C})$. On note $\lambda \in \mathbb{C}$ une valeur propre de S et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{C}) - \{0\}$ tel que $SY = \lambda Y$.
On fixe un indice $i \in [\![1, n]\!]$ tel que $|y_i|$ est maximal ; quitte à diviser Y par y_i, on peut supposer $y_i = 1$.

3. Montrer les égalités et inégalités suivantes :

$$|\lambda| = \left| \sum_{j=1}^{n} S_{i,j} y_j \right| \leq \sum_{j=1}^{n} S_{i,j} |y_j| \leq \sum_{j=1}^{n} S_{i,j} = 1.$$

4. En déduire que $|\lambda| \leq 1$. En étudiant les cas d'égalité dans les inégalités précédentes, montrer que si $|\lambda| = 1$, alors $\lambda = 1$ et $Y \in \mathrm{Vect}(X)$.

On suppose désormais que S est ergodique. On fixe $p \in \mathbb{N}^*$ tel que S^p est strictement positive. On note $\lambda \in \mathbb{C}$ une valeur propre de S et $Y \in \mathcal{M}_{n,1}(\mathbb{C}) - \{0\}$ tel que $SY = \lambda Y$.

[1] Oskar Perron, 1880-1975
[2] Ferdinand Georg Frobenius, 1849-1917

5. Montrer que S^p et S^{p+1} sont strictement positives.
   En déduire que $|\lambda| \leq 1$ et que, si $|\lambda| = 1$, alors $\lambda = 1$ et $Y \in \mathrm{Vect}(X)$.

On note $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ les valeurs propres de S différentes de 1. On admet[3] que S est semblable à une matrice diagonale par blocs de la forme

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

où les matrices A_i sont triangulaires supérieures avec une diagonale de λ_i.

6. Soit $\lambda \in \mathbb{C}$ tel que $|\lambda| < 1$, soit $T \in \mathscr{M}_n(\mathbb{C})$ une matrice triangulaire supérieure dont les coefficients diagonaux sont égaux à λ. En décomposant T sous la forme $\lambda I_n + N$, où N est triangulaire supérieure stricte, montrer que T^p converge vers la matrice nulle quand $p \to +\infty$.
7. En déduire la limite de la suite $(S'^p)_{p\in\mathbb{N}}$.
8. En déduire que $(S^p)_{p\in\mathbb{N}}$ converge vers une matrice, qu'on note S_∞.
9. Montrer que S_∞ est stochastique et de rang 1.
10. En déduire qu'il existe $L = \begin{pmatrix} p_1 & p_2 & \dots & p_n \end{pmatrix} \in \mathscr{M}_{1,n}(\mathbb{R})$ telle que $\forall j \in [\![1,n]\!], p_j \geq 0, \sum_{j=1}^{n} p_j = 1$ et $S_\infty = XL$.
11. Montrer que $LS = L$ et que si $\widetilde{L} \in \mathscr{M}_{1,n}(\mathbb{R})$ est telle que $\widetilde{L}S = \widetilde{L}$, alors $\widetilde{L} \in \mathrm{Vect}(L)$.

2 Théorème ergodique sur les chaînes de Markov

On considère un graphe dont les sommets sont numérotés de 1 à n. On se déplace, en tout temps $p \in \mathbb{N}$, de sommet en sommet ; la décision de rejoindre tel ou tel sommet (ou de ne pas bouger) est aléatoire mais ne dépend que du sommet sur lequel on se situe au moment de prendre la décision.

Formellement, une *chaîne de Markov*[4] (homogène) $\mathscr{M}$ à valeurs dans $[\![1,n]\!]$ est une suite $(X_p)_{p\in\mathbb{N}}$ de variables aléatoires à valeurs dans $[\![1,n]\!]$ telle que la propriété de Markov suivante est vérifiée :

$$\forall p \in \mathbb{N}, \forall i_0, \dots, i_{p+1} \in [\![1,n]\!], P\big((X_0 = i_0) \cap \dots \cap (X_p = i_p)\big) > 0 \implies$$
$$P\big(X_{p+1} = i_{p+1} \mid (X_0 = i_0) \cap \dots \cap (X_p = i_p)\big) = P\big(X_{p+1} = i_{p+1} \mid X_p = i_p\big).$$

On demande de plus que le membre de droite ne dépende pas du temps $p \in \mathbb{N}$.
On note $P \in \mathscr{M}_n(\mathbb{R})$ la *matrice de transition* de $\mathscr{M}$, de coefficients $P(i,j) = P(X_1 = j \mid X_0 = i)$.[5]

Si X est une variable aléatoire à valeurs dans $[\![1,n]\!]$, on note la loi de X par la matrice ligne $\pi_X = \begin{pmatrix} \pi_X(1) & \dots & \pi_X(n) \end{pmatrix} \in \mathscr{M}_{1,n}(\mathbb{R})$, avec $\pi_X(i) = P(X = i)$, pour tout $i \in [\![1,n]\!]$. Pour tout $j \in \mathbb{N}$, on abrège π_{X_j} en π_j et de même, pour tout $i \in [\![1,n]\!]$, $\pi_{X_j}(i)$ en $\pi_j(i)$.

[3] Conséquence de la réduction des endomorphismes, programme de 2ème année.
[4] Andreï Markov, 1856-1922
[5] On suppose pour simplier que $P(X_0 = i) > 0$, pour tout i.

12. Montrer que, pour tout $p \in \mathbb{N}$ et pour tous $i_0, \dots, i_p \in [\![1, n]\!]$,

$$P\big((X_0 = i_0) \cap \cdots \cap (X_p = i_p)\big) = \pi_0(i_0) P(i_0, i_1) P(i_1, i_2) \dots P(i_{p-1}, i_p).$$

13. Montrer que la matrice P est stochastique.

14. Montrer que pour tout $p \in \mathbb{N}$, $\pi_{p+1} = \pi_p P$.

Une mesure de probabilité sur $[\![1, n]\!]$, représentée par une matrice ligne $\nu \in \mathscr{M}_{1,n}(\mathbb{R})$, est *invariante* pour $\mathscr{M}$ si $\pi_0 = \nu \implies \forall p \in \mathbb{N}, \pi_p = \nu$.

15. Montrer que ν est invariante pour $\mathscr{M}$ ssi $\nu P = \nu$.

La chaîne de Markov $\mathscr{M}$ est *fortement irréductible* s'il existe un entier p tel que, pour tous i, j dans $[\![1, n]\!]$, $P(X_p = j \mid X_0 = i) > 0$.

16. Montrer que $\mathscr{M}$ est fortement irréductible ssi sa matrice de transition P est ergodique.

17. En déduire que si $\mathscr{M}$ est fortement irréductible, alors il existe une unique mesure de probabilité ν invariante pour $\mathscr{M}$ sur $[\![1, n]\!]$. Montrer de plus que $\pi_p \to \nu$, quand $p \to +\infty$[6].

3 Deux exemples

3.1 Chaîne de Markov à deux états

On considère une chaîne de Markov avec seulement deux sommets 1 et 2. La probabilité de passer du sommet 1 au sommet 2 est notée $\alpha \in]0, 1[$, celle de passer du sommet 2 au sommet 1 est notée $\beta \in]0, 1[$.

18. Représenter la situation sous la forme d'un graphe et déterminer la matrice de transition Q.

19. Montrer que la chaîne de Markov est fortement irréductible. Déterminer la loi invariante ν, puis la limite de Q^p, quand $p \to +\infty$.

3.2 Urnes d'Ehrenfest

On considère deux urnes U_1 et U_2 contenant au total n boules. A chaque instant $p \in \mathbb{N}$, on tire au hasard une boule et on la place dans l'autre urne. On note X_p le nombre de boules dans l'urne 1 au temps p (avant l'échange).

20. Modéliser la situation comme une chaîne de Markov dont l'ensemble des états est $[\![0, n]\!]$[7]. La représenter sous la forme d'un graphe.

21. Écrire la matrice de transition correspondante.

22. La chaîne de Markov est-elle fortement irréductible ?

23. Montrer que la loi binomiale $\mathscr{B}(n, 1/2)$ est l'unique loi invariante.

24. Y a-t-il en général convergence de la loi de X_p vers la loi $\mathscr{B}(n, 1/2)$, quand $p \to +\infty$ (au sens de la question 17) ?

---

[6] la convergence ayant lieu coefficient par coefficient.
[7] et non plus $[\![1, n]\!]$ comme précédemment.