

Cours d'analyse de données en géographie

Niveau Master 1 - GEANDO

Pour commencer avec de bonnes bases...

Statistiques

Maxime Forriez^{1,a}

¹ Institut de géographie, 191, rue Saint-Jacques, Bureau 105, 75 005 Paris,
^amaxime.forriez@sorbonne-universite.fr

16 septembre 2025

Exercice 1

Soit la matrice stochastique régulière $P = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

1. Chercher un vecteur de probabilité à deux composantes, noté $t = \begin{pmatrix} x \\ 1-x \end{pmatrix}$ tel que $t.P = t$.
2. Vers quelle matrice T converge la suite P, P^2, P^3, \dots ?

Question 1. Méthode 1

$$t.P = t \Leftrightarrow \begin{pmatrix} x & 1-x \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} x & 1-x \end{pmatrix} \quad (1)$$

$$\Leftrightarrow \left[\begin{pmatrix} x & 1-x \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \quad \begin{pmatrix} x & 1-x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} \right] = \begin{pmatrix} x & 1-x \end{pmatrix} \quad (2)$$

$$\Leftrightarrow \left[0x + \frac{1}{2}(1-x) \quad x + \frac{1}{2}(1-x) \right] = \begin{pmatrix} x & 1-x \end{pmatrix} \quad (3)$$

$$\Leftrightarrow \begin{pmatrix} \frac{1}{2} - \frac{1}{2}x & x + \frac{1}{2} - \frac{1}{2}x \end{pmatrix} = \begin{pmatrix} x & 1-x \end{pmatrix} \quad (4)$$

$$\Leftrightarrow \begin{cases} x = \frac{1}{2} - \frac{1}{2}x \\ \frac{1}{2} + \frac{1}{2}x = 1 - x \end{cases} \quad (5)$$

$$\Leftrightarrow \begin{cases} x + \frac{1}{2} = \frac{1}{2} \\ x + \frac{1}{2}x = 1 - \frac{1}{2} = \frac{1}{2} \end{cases} \quad (6)$$

$$\Leftrightarrow \begin{cases} x = \frac{1}{3} \\ x = \frac{1}{3} \end{cases} \quad (7)$$

On en déduit que $t = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}$. Il constitue l'unique vecteur de probabilité constant de P .

Question 2. Méthode 2

On pose $u = (x, y)$. On cherche à résoudre :

$$(x, y) \cdot \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (x, y) \quad (8)$$

$$\Leftrightarrow \left[(x \ y) \cdot \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \quad (x \ y) \cdot \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} \right] = (x \ y) \quad (9)$$

$$\Leftrightarrow \left[\frac{1}{2}y \quad x + \frac{1}{2}y \right] = (x \ y) \quad (10)$$

$$\Leftrightarrow \begin{cases} x = \frac{1}{2}y \\ x + \frac{1}{2}y = y \end{cases} \quad (11)$$

$$\Leftrightarrow \begin{cases} x = \frac{1}{2}y \\ y = \frac{1}{2}y + \frac{1}{2}y \end{cases} \quad (12)$$

$$\Leftrightarrow \begin{cases} x = \frac{1}{2}y \\ y = y \end{cases} \quad (13)$$

$$\Leftrightarrow \begin{cases} y = 2x \\ y - \frac{1}{2}y = x \end{cases} \quad (14)$$

$$\Leftrightarrow \begin{cases} y = 2x \\ \frac{1}{2}y = x \end{cases} \quad (15)$$

$$\Leftrightarrow \begin{cases} y = 2x \\ y = 2x \end{cases} \quad (16)$$

Si $x = 1$, alors $y = 2$. On en déduit $u = (1, 2)$. Pour obtenir un vecteur de probabilité, on divise par le total de la ligne, soit 3, donc $u = \left(\frac{1}{3}, \frac{2}{3}\right)$.

On en déduit que :

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \quad (17)$$

Question 2

La suite P, P^2, P^3, \dots converge vers la matrice T dont les lignes sont toutes égales au vecteur t .

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \approx \begin{pmatrix} 0,33 & 0,67 \\ 0,33 & 0,67 \end{pmatrix} \quad (18)$$

$$P^2 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix} \approx \begin{pmatrix} 0,25 & 0,75 \\ 0,25 & 0,75 \end{pmatrix} \quad (19)$$

$$P^3 = \begin{pmatrix} \frac{1}{8} & \frac{3}{8} \\ \frac{1}{8} & \frac{3}{8} \end{pmatrix} \approx \begin{pmatrix} 0,125 & 0,875 \\ 0,125 & 0,875 \end{pmatrix} \quad (20)$$

$$P^4 = \begin{pmatrix} \frac{3}{16} & \frac{5}{16} \\ \frac{3}{16} & \frac{5}{16} \end{pmatrix} \approx \begin{pmatrix} 0,1875 & 0,8125 \\ 0,1875 & 0,8125 \end{pmatrix} \quad (21)$$

$$P^5 = \begin{pmatrix} \frac{5}{32} & \frac{11}{32} \\ \frac{5}{32} & \frac{11}{32} \end{pmatrix} \approx \begin{pmatrix} 0,15625 & 0,9375 \\ 0,15625 & 0,9375 \end{pmatrix} \rightarrow T \quad (22)$$

Exercice 2

Déterminer l'unique vecteur de probabilité constant de la matrice stochastique régulière :

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \quad (23)$$

Soit $u = (x \ y \ z)$ un vecteur de probabilité constant.

$$(x \ y \ z) \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} = (x \ y \ z) \quad (24)$$

$$\Leftrightarrow \left[(x \ y \ z) \cdot \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix} \quad (x \ y \ z) \cdot \begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix} \quad (x \ y \ z) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right] = (x \ y \ z) \quad (25)$$

$$\Leftrightarrow (0x + 0y + \frac{1}{2}z \quad x + 0y + \frac{1}{2}z \quad 0x + y + 0z) = (x \ y \ z) \quad (26)$$

$$\Leftrightarrow (\frac{1}{2}z \quad x + \frac{1}{2}z \quad y) = (x \ y \ z) \quad (27)$$

$$\Leftrightarrow \begin{cases} \frac{1}{2}z = x \\ x + \frac{1}{2}z = y \\ y = z \end{cases} \quad (28)$$

On obtient un système homogène avec une infinité de solution.

On choisit $z = 1 \Leftrightarrow \begin{cases} x = \frac{1}{2} \\ y = 1 \end{cases}$ Pour obtenir un vecteur de probabilité, il suffit de multiplier u par l'inverse de $\frac{1}{2} + 1 + 1 = \frac{5}{2}$:

$$\frac{2}{5} \begin{pmatrix} \frac{1}{2} & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{pmatrix} \quad (29)$$

Pour éviter le calcul fractionnel, il suffit de prendre $z = 2 \Leftrightarrow \begin{cases} x = 1 \\ y = 2 \end{cases}$ $x + y + z = 1 + 2 + 2 = 5$, soit :

$$\frac{1}{5} \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{pmatrix} \quad (30)$$

On obtient ainsi un vecteur de probabilité constant.

Exercice 3

Un ouvrier se rend chaque matin à son travail soit en voiture, soit en train. On suppose qu'il ne prend jamais le train deux jours consécutifs, mais s'il se rend à son travail en voiture, il peut le jour suivant tout aussi bien prendre sa voiture, ou prendre le train.

1. Définir l'espace des états du système. On pose v pour la voiture et t pour le train.
 - On pose v pour la voiture et t pour le train, soit $\{t, u\}$.
2. Justifier pourquoi il s'agit d'une chaîne de Markov.
 - Le processus stochastique est une chaîne de Markov puisque le résultat du jour donné dépend **uniquement** de ce qui s'est passé le jour précédent.
3. Établir la matrice de transition.

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad (31)$$

- Les lignes correspondent à la situation de la veille. Les colonnes correspondent aux modes de transport possibles du jour.
 - La ligne 1 représente le train. Si la veille, l'ouvrier a pris le train, la probabilité qu'il prenne le train le lendemain est nulle. Si la veille, l'ouvrier a pris sa voiture, la probabilité qu'il prenne sa voiture le lendemain est certaine.
 - La ligne 2 représente la voiture. Si la veille, l'ouvrier a pris sa voiture, la probabilité qu'il prenne le train le lendemain est $\frac{1}{2}$. Si la veille, l'ouvrier a pris sa voiture la probabilité qu'il prenne sa voiture le lendemain est $\frac{1}{2}$.
4. Quelle est la matrice de transition matérialisant le choix du mode de transport de l'ouvrier le quatrième jour?
 - Le quatrième jour, la matrice de transition matérialisant le choix du mode de transport de l'ouvrier est P^4 .

$$P^4 = \begin{pmatrix} \frac{3}{32} & \frac{5}{32} \\ \frac{8}{32} & \frac{21}{32} \end{pmatrix} \quad (32)$$

- La probabilité pour que le système évolue de l'état t à l'état v en quatre transitions **exactement** est $\frac{5}{8}$.
 - La probabilité pour que le système évolue de l'état t à l'état t en quatre transitions **exactement** est $\frac{3}{8}$.
 - La probabilité pour que le système évolue de l'état v à l'état t en quatre transitions **exactement** est $\frac{11}{32}$.
 - La probabilité pour que le système évolue de l'état v à l'état v en quatre transitions **exactement** est $\frac{21}{32}$.
5. Si le premier jour, l'ouvrier jette un dé bien équilibré et décide de ne partir travailler en voiture que si et seulement s'il obtient un 6, calculer l'état dans quatre jours.
- En d'autres termes, $p^{(0)} = \left(\frac{5}{6} \quad \frac{1}{6} \right)$ est la distribution de probabilité initiale. L'état dans quatre jours est : $p^{(4)}$
- $$p^{(4)} = p^{(0)} P^4 \quad (33)$$
- $$p^{(4)} = \left(\frac{5}{6} \quad \frac{1}{6} \right) \cdot \begin{pmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{11}{32} & \frac{21}{32} \end{pmatrix} = \left(\frac{35}{96} \quad \frac{61}{96} \right) \quad (34)$$
- La probabilité que l'ouvrier prenne le train dans quatre jours **exactement** est $\frac{35}{96}$.
 - La probabilité que l'ouvrier prenne sa voiture dans quatre jours **exactement** est $\frac{61}{96}$.
6. Calculer l'état stationnaire à n itérations.
- L'état stationnaire des choix de transport de l'ouvrier converge vers le vecteur $\left(\frac{1}{3} \quad \frac{2}{3} \right)$. À n itérations, l'ouvrier a $\frac{1}{3}$ de chances de choisir le train, et $\frac{2}{3}$ sa voiture.

Exercice 4

Trois garçons A, B, C jouent au ballon.

- A lance **toujours** le ballon à B.
 - B lance **toujours** le ballon à C.
 - C peut lancer **tout aussi bien** le ballon à A ou B.
1. Quel est l'espace des états ?
 - L'espace des états est $\{A, B, C\}$.
 2. Justifier pourquoi il s'agit d'une chaîne de Markov.
 - C'est une chaîne de Markov puisque le joueur qui jette le ballon n'est pas influencé par ceux qui ont eu le ballon avant lui.
 3. Calculer la matrice de transition.

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \quad (35)$$

N.B. Il est impossible de se lancer le ballon à soi-même. La diagonale est nulle.

4. Si C est le premier à avoir le ballon, quelle est la probabilité pour que A ait le ballon après trois lancers ?

— Si C est le premier à avoir le ballon, cela signifie que $p^{(0)} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ est la distribution initiale alors :

$$p^{(1)} = p^{(0)}.P = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \quad (36)$$

$$p^{(2)} = p^{(1)}.P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad (37)$$

$$p^{(3)} = p^{(2)}.P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} .P = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \quad (38)$$

— Après trois lancers de ballon, la probabilité pour que A ait le ballon est $\frac{1}{4}$; B, $\frac{1}{4}$ et C, $\frac{1}{2}$.

— Si on veut directement le résultat, il suffit de calculer :

$$p^{(3)} = p^{(0)}.P^3 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \quad (39)$$

— Si A est le premier à avoir le ballon, alors $p^{(0)} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$, donc $p^{(3)} = p^{(0)}.P^3 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

— Si B est le premier à avoir le ballon, alors $p^{(0)} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$, donc $p^{(3)} = p^{(0)}.P^3 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

5. Calculer l'état stationnaire à n itérations.

— L'état stationnaire des passes est le vecteur $\begin{pmatrix} \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{pmatrix}$. À n itérations, A a $\frac{1}{5}$ de chances d'avoir le ballon ; B, $\frac{2}{5}$; C, $\frac{2}{5}$.