Rapport TAI (Chaînes de Markov)

Azad BADAKOL, Guillaume DELHAYE Romain, DROUET Antoine IA, Djamel NAFISSE

Sommaire

1. La forêt des Landes : modèle à 3 états

1) Comment se traduisent ces informations sur la matrice P telle que nous l'avons définie, possédant la propriété de Markov ?

$$X_n = \begin{pmatrix} Rs \\ Ch \\ Cht \end{pmatrix}_n = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$X_{n+1} = \begin{pmatrix} Rs \\ Ch \\ Cht \end{pmatrix}_{n+1} = \begin{pmatrix} 0.5 \\ 0.2 \\ 0.3 \end{pmatrix}$$

$\{a, b, c, d, e, f\} \in R^6$

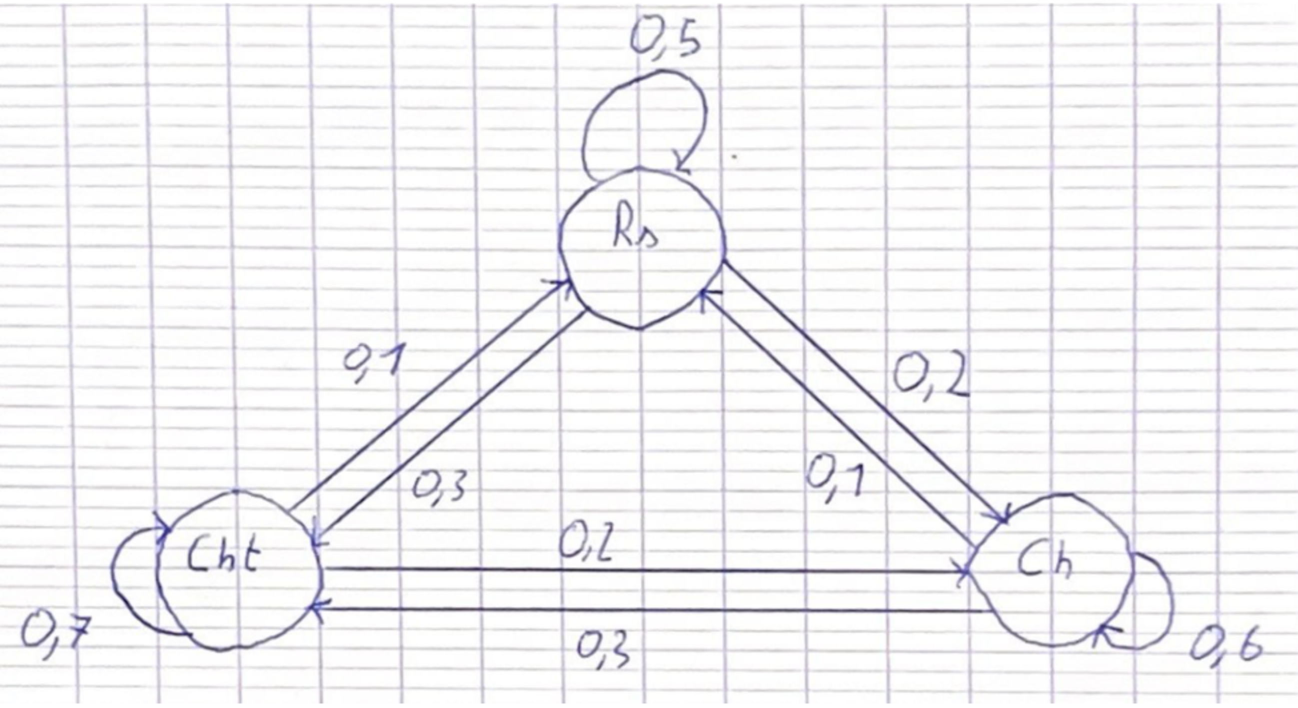
$$P \times X_n = \begin{pmatrix} 0.5 & a & b \\ 0.2 & c & d \\ 0.3 & e & f \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.2 \\ 0.3 \end{pmatrix} = X_{n+1}$$

Donc $P = \begin{pmatrix} 0.5 & a & b \\ 0.2 & c & d \\ 0.3 & e & f \end{pmatrix}$

---

2) Dessiner un graphe qui modélise l'écosystème. On précisera notamment les valeurs sur chacun des arcs.

$$P = \begin{pmatrix} 0.5 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.3 & 0.7 \end{pmatrix}$$



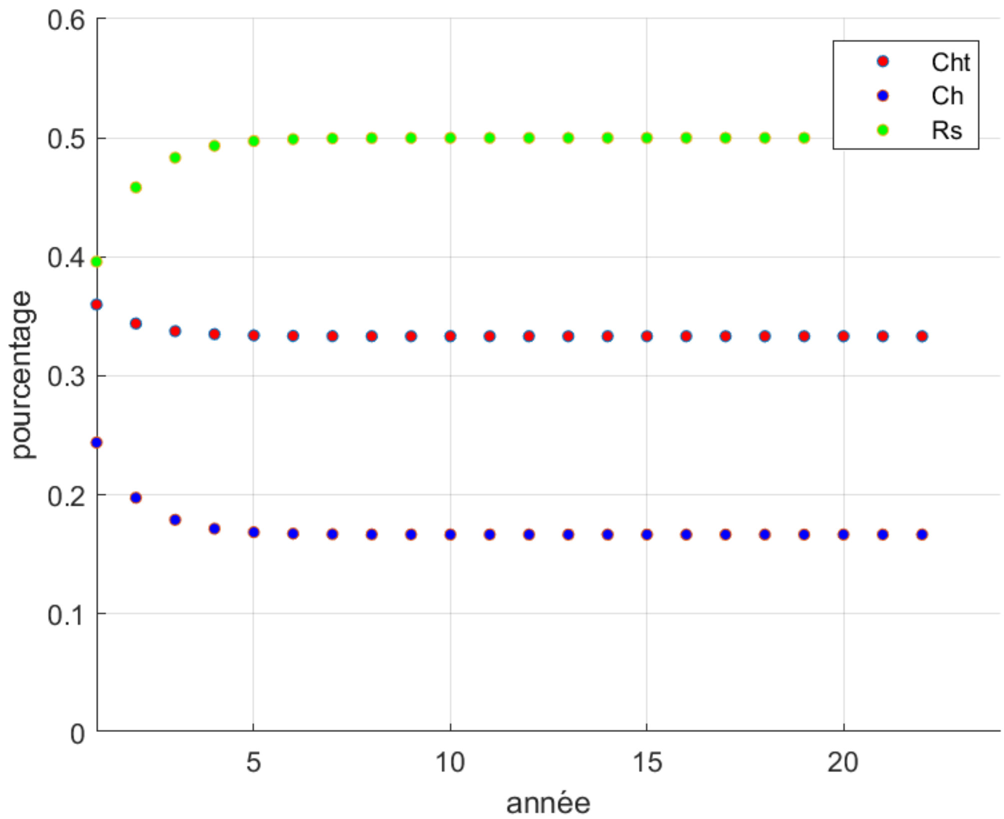
---

3) Calculez la proportion de chaque arbre pour l'année n = 1.

$$X_1 = P \times X_0 = \begin{pmatrix} 0.5 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.3 & 0.7 \end{pmatrix} \times \begin{pmatrix} 0.36 \\ 0.40 \\ 0.24 \end{pmatrix} = \begin{pmatrix} 0.24 \\ 0.36 \\ 0.40 \end{pmatrix}$$

---

4 & 5) A l'aide d'un ordinateur, calculez la proportion de chaque arbre pour n = 20.



On peut donc en déduire que l'écosystème tend vers la stabilité. Précisément, il tend vers cette valeur :

$$X_{20} = \begin{pmatrix} 0.50 \\ 0.33 \\ 0.16 \end{pmatrix}$$

---

6) Diagonalisation de P :

$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2/5 & 0 \\ 0 & 0 & 2/5 \end{pmatrix}$ Chaque élément non nul est une valeur propre de D

$A = \begin{pmatrix} 1/3 & -1 & -1 \\ 2/3 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ La première colonne est le vecteur propre de la première valeur de D et ainsi de suite.

$$A^{-1} = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ -1/3 & 2/3 & -1/3 \\ -1/2 & -1/2 & 1/2 \end{pmatrix}$$

$$P = A \times D \times A^{-1}$$

---

7) Écrire le vecteur X_n en fonction de D^n, A, A^{-1} et la distribution initiale X_0

$$X_n = X_0 \times (A \times D^n \times A^{-1})^T$$

Avec matlab nous obtenons l'exact même résultat qu'à l'exercice 5.

2. Ecosystèmes méditerranéens à 5 états

1) Dessiner le graphe relatif à ce modèle.

2) Quelle est, selon ce modèle, la probabilité que l'on passe de l'état Pi à Pe, d'une année sur l'autre ?

$$Pi - Pe = 0.25$$

Et de l'état P_e à P_i ?
$Pe - Pi = 0$ car en une année, Pe ne peut que passer à Ga ou rester à Pe.

---

3) Sachant que l'on est dans l'état Pi, quelle est la probabilité d'une trajectoire du type $Pi - Pe - Ga$?

$$P((Pe \cap Ga) \setminus Pi) = P(Pe \setminus Pi) \times P(Ga \setminus Pe) = 0.25 \times 0.6 = 0.15$$

Et d'une trajectoire du type $Pi - C - V$?
$P((C \cap V) \setminus Pi) = P(C \setminus Pi) * P(V \setminus C) = 0.2 \times 0.1 = 0.02$

---

4) Donnez un exemple de trajectoire de probabilité nulle.

$P_i - V - G_a$ est un exemple de trajectoire de probabilité nulle

---

5) En vous servant de votre cours d'Algèbre linéaire de L1 expliquez, lorsque n tend vers l'infini, vers quelle proportion tend la dynamique lorsque $X_0 = \begin{pmatrix} C \\ V \\ Pe \\ Ga \\ Pi \end{pmatrix}_0 = 0.2 \times \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

$$X_0 = \begin{pmatrix} C \\ V \\ Pe \\ Ga \\ Pi \end{pmatrix}_0 = 0.2 \times \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$X_1 = P \times X_0$, alors

$$X_n = P^n \times X_0 = \begin{pmatrix} 0.1752 \\ 0.1168 \\ 0.2044 \\ 0.1533 \\ 0.3504 \end{pmatrix}$$

---

6) La transition de P_i vers P_e mime la présence des incendies. Étudier l'influence aux temps longs des incendies, en faisant varier sa valeur dans la matrice.

La valeur initiale est $P_i - P_e = 0.25$ et pour cette valeur : $X_n = \begin{pmatrix} 0.1752 \\ 0.1168 \\ 0.2044 \\ 0.1533 \\ 0.3504 \end{pmatrix}$

Si $Pi - Pe = 0$ alors $X_n = 10^{-6} \times \begin{pmatrix} 0.7431 \\ 0.6639 \\ 0.3802 \\ 0.3151 \\ 0.9205 \end{pmatrix}$

Si $Pi - Pe = 0.1$, $X_n = 10^{-3} \times \begin{pmatrix} 0.0668 \\ 0.0533 \\ 0.0473 \\ 0.0378 \\ 0.1006 \end{pmatrix}$

Si $Pi - Pe = 0.4$, $X_n = 10^{3} \times \begin{pmatrix} 0.4955 \\ 0.2809 \\ 0.8967 \\ 0.6309 \\ 1.2529 \end{pmatrix}$

Si $Pi - Pe = 0.6$, $X_n = 10^{7} \times \begin{pmatrix} 0.5804 \\ 0.2772 \\ 1.6602 \\ 1.0842 \\ 1.8503 \end{pmatrix}$

Si $Pi - Pe = 0.8$, $X_n = 10^{10} \times \begin{pmatrix} 1.4725 \\ 0.6174 \\ 5.9537 \\ 3.6564 \\ 5.5508 \end{pmatrix}$

On remarque que X_n devient plus grand quand $P_i - P_e$ devient plus grand. On peut en déduire que les incendies entraînent plus d'arbres sur le long terme et à l'inverse, pas d'incendies réduit le nombre d'arbres sur le long terme.

3. Modèle plus riche d'écosystème

$$P = \begin{pmatrix}
0.1 & 0 & 0.8 & 0 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.2 & 0.4 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.1 & 0.2 & 0.05 & 0.9 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5 & 0.2 & 0.05 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.1 & 0.2 & 0.05 & 0.05 & 0.7 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.6 & 0.4 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.2 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0 & 0.3 & 0.3 & 0.3 & 0.4 & 0.1 & 0.3 & 0 & 0.15 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.05 & 0.3 & 0.1 & 0.1 & 0.1 & 0.2 & 0.05 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.1 & 0 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.15 & 0.05 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.05 & 0.05 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.1 & 0 & 0.1 & 0.2 & 0.1 & 0.2 & 0.3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0.1 & 0.05 & 0.2 & 0.1 & 0.1 & 0.2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0.05 & 0.1 & 0.05 & 0.1 & 0.1 & 0.2 & 0.1
\end{pmatrix}$$

---

1) A partir de chaque état, le ou lesquels sont à un moment ou un autre visités ? Donner alors les groupes d'états qui communiquent tous ensemble. Expliquer comment cela se voit sur la matrice.

Tous sont visités sauf le 12e élément. Le 1er au 4e communiquent ensemble, le 6,7 et 8e, communiquent entre eux. Les 13 à 20 aussi. La 5e ligne communique avec les 1,3,5 et est visitee par le 7e element. Les 9,10 et 11 communquent entre elles, la 10 visite la 20 et la 11 visite la 13. Il faut regarder les lignes pour regarder quel élément est visité et les colonnes pour regarder les éléments que visite celui sur lequel on est.

---

2) Calculs de X_1, X_2, X_{50} avec matlab :

Code matlab utilisé dans le reste de l'exercice :

```
P = [0.1 0 0.8 0 0.2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
    0.2 0.4 0.05 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
    0.1 0.2 0.05 0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
    0.5 0.2 0.05 0.05 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
    0.1 0.2 0.05 0.05 0.7 0 0 0 0.2 0 0 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 0 0 0.6 0.4 0.1 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 0 0 0.1 0.2 0.7 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 0 0 0.1 0.2 0.1 0 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0;
    0 0 0 0 0 0 0 0 0 0 0.1 0 0.3 0.3 0.3 0.4 0.1 0.3 0 0.15;
    0 0 0 0 0 0 0 0 0 0 0 0 0.3 0.05 0.3 0.1 0.1 0.1 0.2 0.05;
    0 0 0 0 0 0 0 0 0 0 0 0 0.1 0.1 0 0.1 0.1 0.1 0.1 0.1;
    0 0 0 0 0 0 0 0 0 0 0 0 0.1 0.1 0.1 0.1 0.1 0.1 0.15 0.05;
    0 0 0 0 0 0 0 0 0 0 0 0 0.1 0.1 0.1 0.1 0.1 0.1 0.05 0.05;
    0 0 0 0 0 0 0 0 0 0 0 0 0.1 0.1 0 0.1 0.2 0.1 0.2 0.3;
    0 0 0 0 0 0 0 0 0 0 0 0 0 0.2 0.1 0.05 0.2 0.1 0.1 0.2;
    0 0 0 0 0 0 0 0 0 0.2 0 0 0 0.05 0.1 0.05 0.1 0.1 0.2 0.1;
];

X0 = 1/16*[1 1 1 1 1 0 0 0 1 1 1 0 1 1 1 1 1 1 1 1]'

result = [X0]

length = 10;
green = [0. 1, 0];
red = [1, 0. 0];
colors_p = [linspace(green(1),red(1),length)', linspace(green(2),red(2),length)', linspace(green(3),red(3),length)'];

for x = 2:50
    result(x,:) = P*result(x-1,:);
end
round(result, 4);

hold on
for x = 1:10
    plot(result(x, :), "gs", "MarkerFaceColor", colors_p(x,:))
end
legend("Pas n=1", "Pas n=2", "Pas n=3", "Pas n=4", "Pas n=5", "Pas n=6", "Pas n=7", "Pas n=8", "Pas n=9", "Pas n=10")

ylabel("Valeurs des états")
xlabel("états 1 à 20")
title("Valeurs des états sur 10 itérations")
hold off
```

$X_0 = 1/16 \times (\ 1\quad 1\quad 1\quad 1\quad 1\quad 0\quad 0\quad 0\quad 1\quad 1\quad 1\quad 0\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\quad 1\ )^T$

$X_1 =$
$(\ 0.0688\quad 0.0406\quad 0.0844\quad 0.0500\quad 0.0813\quad 0\quad 0\quad 0\quad 0.0688\quad 0.0625\quad 0.0250\quad 0\quad 0.1219\quad 0.0750\quad 0.0438\quad 0.0500\quad 0.0438\quad 0.0688\quad 0.($

$X_2 =$
$(\ 0.09060\quad 0.03420\quad 0.07230\quad 0.04920\quad 0.09230\quad 0\quad 0\quad 0\quad 0.06880\quad 0.03690\quad 0.02190\quad 0\quad 0.1281\quad 0.08440\quad 0.04750\quad 0.05200\quad 0.0461$

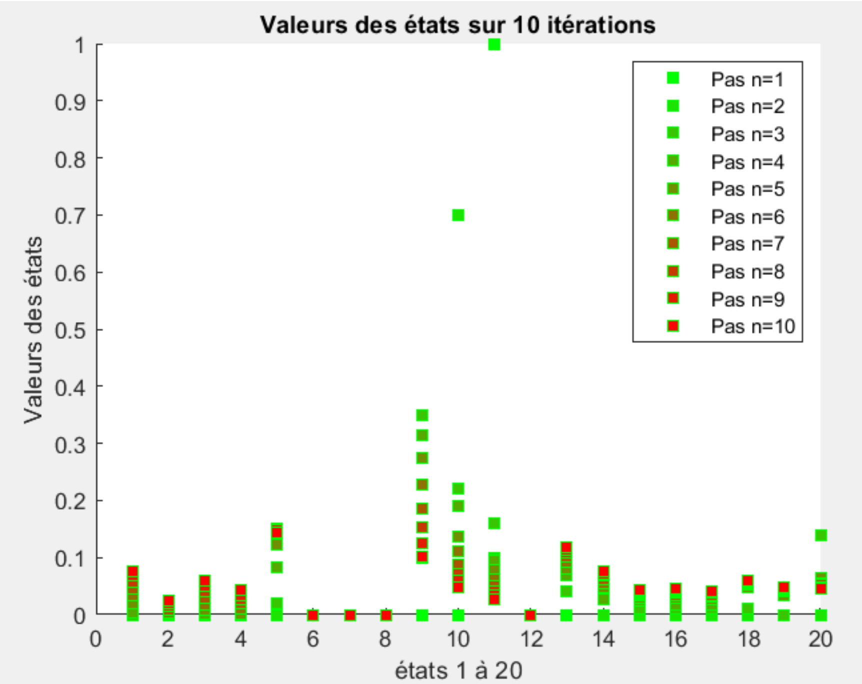
$X_{50} =$
$(\ 0.1084\quad 0.04440\quad 0.09850\quad 0.07160\quad 0.09410\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0.1459\quad 0.09710\quad 0.05300\quad 0.05920\quad 0.05320\quad 0.07270\quad 0.0600$

Il faut en moyenne un seul pas pour que toutes les valeurs soient concentrées sur les états {1, 2, 3, 4, 5}, {13, 14, 15, 16, 17, 18, 19, 20}

---

3) Simulation des états de Y_n par rapport à n :

$Y_0 = (\ 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 1\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\ )^T$



Nombre de pas moyen pour que le système peuple les états {1, 2, 3, 4, 5} : n=4

Nombre de pas moyen pour peupler que le système états {14, 15, 16, 17, 18, 19, 20} : n=3

L'état {1, 2, 3, 4, 5} est le plus peuplé lorsque la dynamique a convergé, pour n=50 :

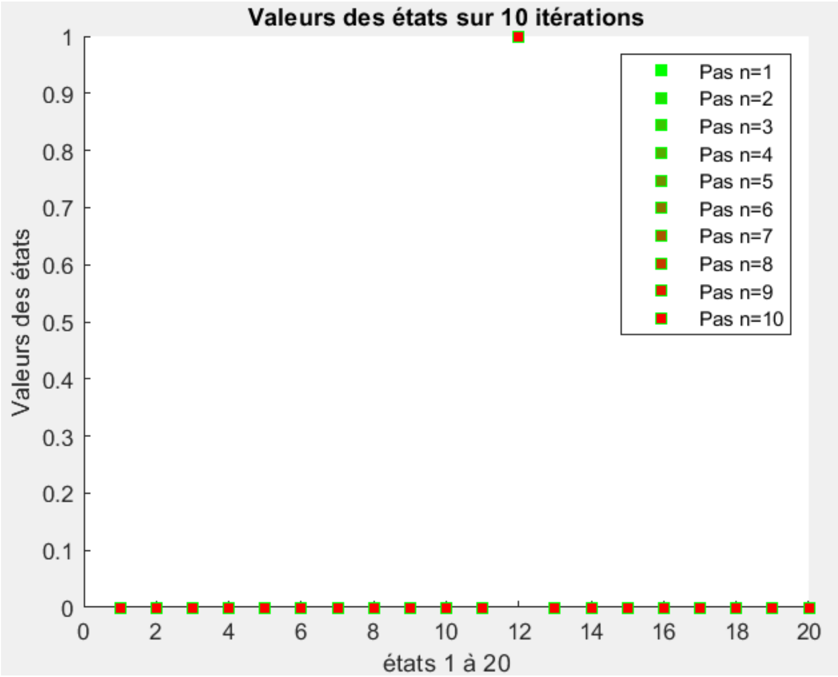
{1, 2, 3, 4, 5} -> {0.1185, 0.0484, 0.1076, 0.0782, 0.1028} -> 0.4555

{14, 15, 16, 17, 18, 19, 20} -> {0.0906, 0.0494, 0.0552, 0.0496, 0.0678, 0.0559, 0.0391} -> 0.4076

---

4) Simulation des états de Z_n par rapport à n :

$Z_0 = (\ 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 1\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\ )^T$

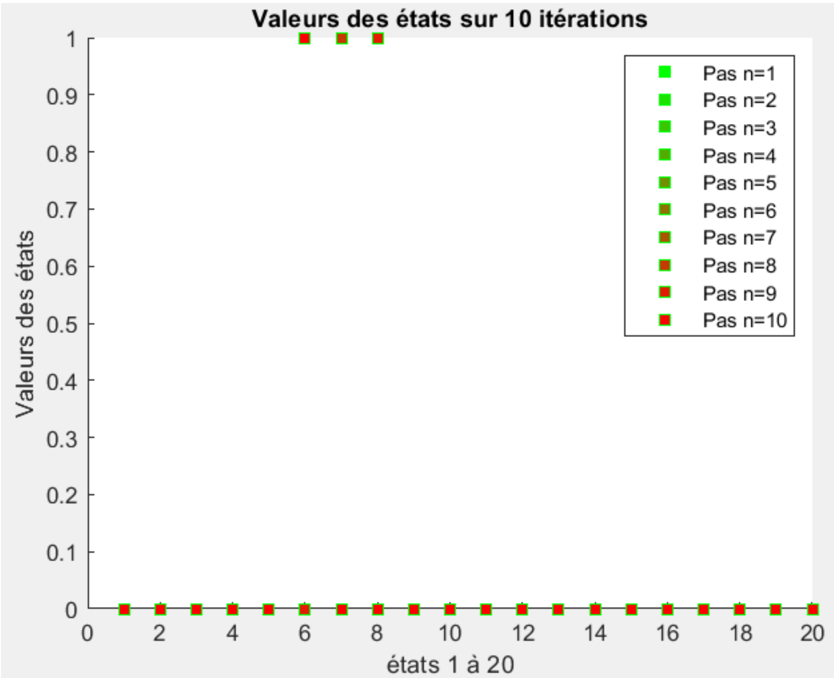


Le 12ème élément de X_0 est le seul élément non nul de ce vecteur.
Ainsi, seuls les éléments de la 12ème colonne de P compte pour le résultat de la multiplication : $X_0 \times P$
Or seul le 12ème éléments de la 12ème colonne de P n'est pas nulle.
Ce qui fait que $X_0 \times P = X_0$ pour chaque itération.

---

5) Simulation des états de A_n par rapport à n :

$$A_0 = (\begin{array}{cccccccccccccccccccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array})^T$$



Seules les états {6, 7, 8} varient, et ce suivant une période n=3 :

	6	7	8
1	1	0	0
2	0	1	0
3	0	0	1
4	1	0	0
5	0	1	0
6	0	0	1

Le système revient à sa distribution initiale de manière périodique, pour tout $n = 3x, x \in N$

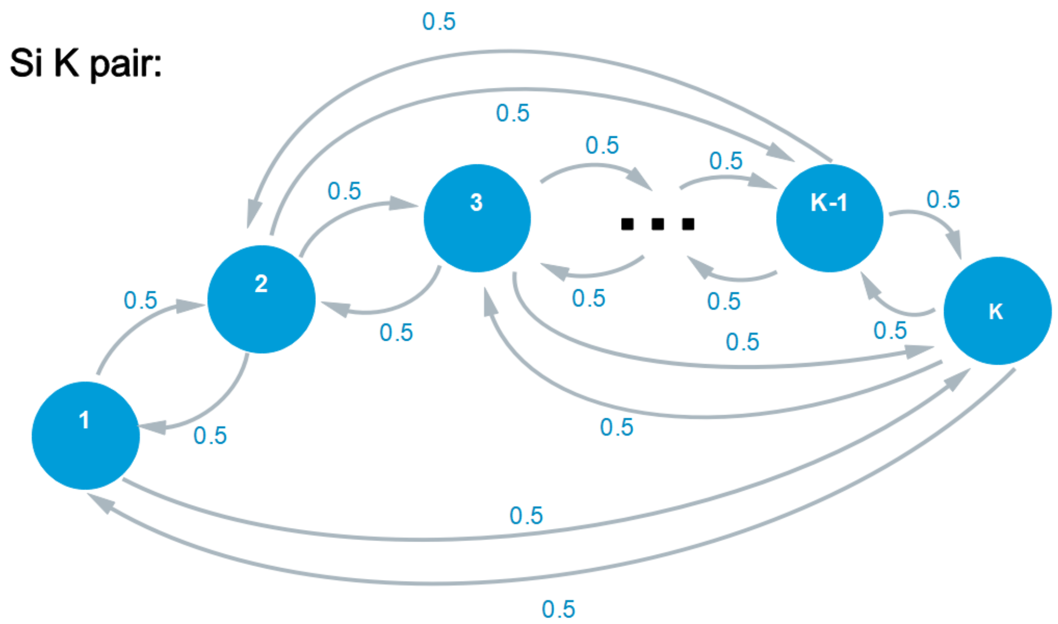
Cette périodicité est dû à la composition des colonnes de P ci dessous :

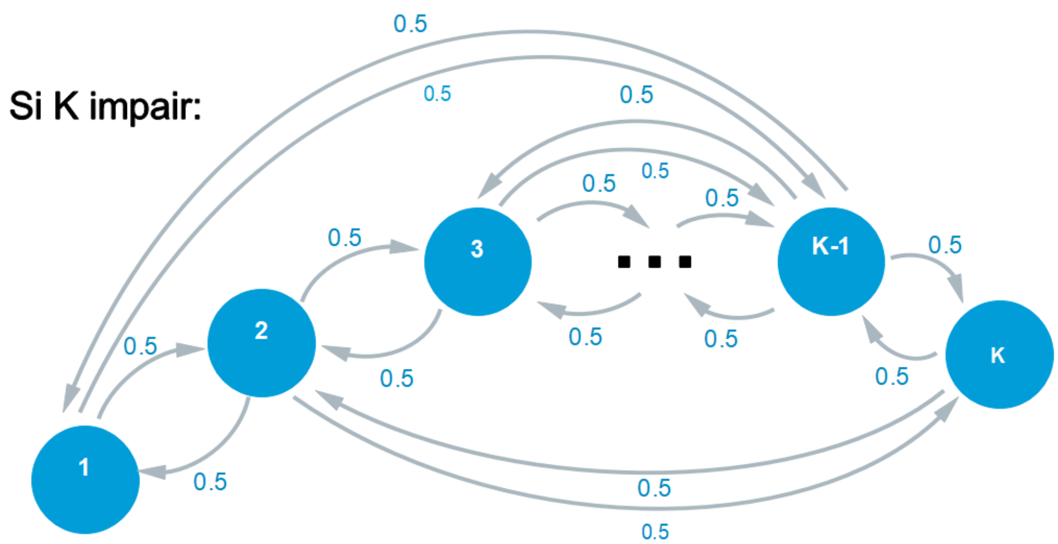
	6	7	8
5	0	0	0
6	0	0	1
7	1	0	0
8	0	1	0

On peut observer que la colonne 6 amène à décaler le 1 du vecteur initial vers l'état 7
La colonne 7 amène à décaler le 1 du vecteur initial vers l'état 8
La colonne 8 amène à décaler le 1 du vecteur initial vers l'état 6
Et la boucle recommence.

4. Pour aller plus loin

1) On remarque que la matrice correspond à une matrice de transition.
On rappelle que le nombre de colonne correspond nombre d'état qu'on appellera k.

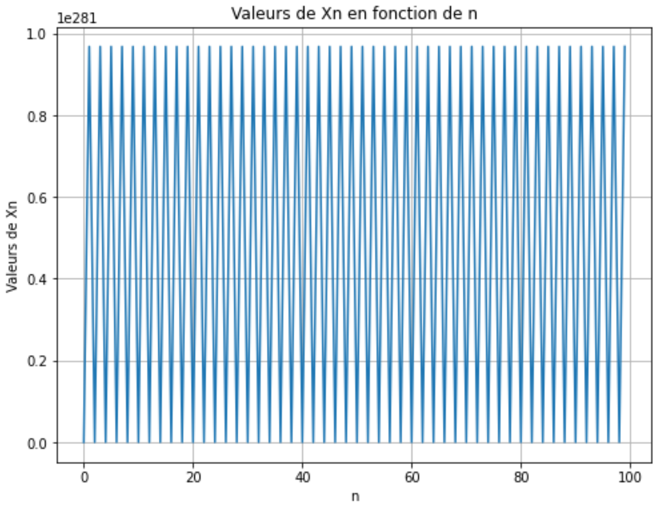


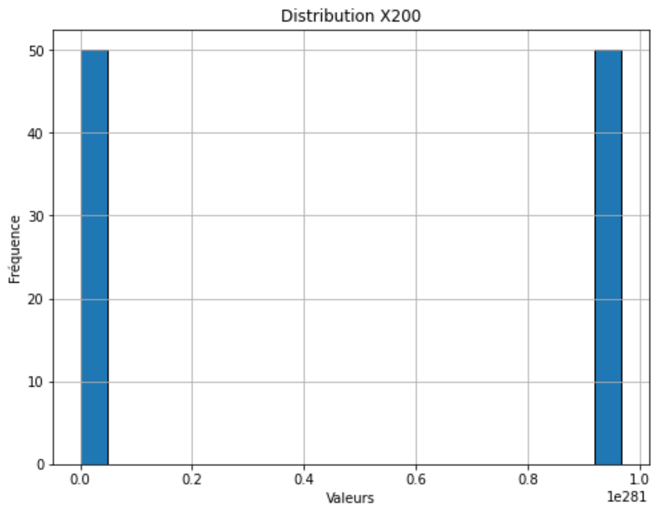


2) On peut expliciter X_n tel que:
$X_n = X_0 \times P^n$

Pour faire la simulation numérique, nous avons utilisé python avec la bibliothèque de mathlab:

```
import numpy as np
import matplotlib.pyplot as plt

# Définir la taille de la chaîne n et le nombre de pas k
k = 100
n = 200

# Matrice de transition P pour une chaîne de Markov
P = np.zeros((k, k)) #on remplit une matrice de 0
#on fait une double boucle pour parcourir la matrice et attribuer les valeurs correspondantes
for i in range(k):
    for j in range(k):
        if i % 2 != 0:
            if j%2==0:
                P[i, j] = 0.5
            else:
                P[i, j] = 0
        elif  i % 2 == 0:
            if j%2!=0:
                P[i, j] = 0.5
            else:
                P[i, j] = 0
        else:
            P[i, j] = 0

#on crée d'abord X0
XO = np.zeros(k) #X0 est un vecteur remplit de 0
#on lui attribut les valeurs correspondantes
for i in range(k):
    if i%2!=0:
        XO[i]=0
    else:
        XO[i]=1
#Nous allons mettre P à la puissance n
Pn = np.copy(P) #on copie P à Pn
#on crée une boucle pour multiplier pn n fois à p
#sa correspond à P puissance n
for a in range(n):
    Pn = np.dot(Pn,P)

# On crée Xn
Xn=np.dot(Pn,XO) #on fait la matrice Pn.X0 et l'attribut à Xn

# Affichage de Xn en fonction de n
plt.figure(figsize=(8, 6))
plt.plot(Xn)
plt.title('Valeurs de Xn en fonction de n')
plt.xlabel('n')
plt.ylabel('Valeurs de Xn')
plt.grid(True)
plt.show()

# Représentation de la distribution X200
plt.figure(figsize=(8, 6))
plt.hist(Xn, bins=20. edgecolor='black')
plt.title('Distribution X200')
plt.xlabel('Valeurs')
plt.ylabel('Fréquence')
plt.grid(True)
plt.show()
```





Les probabilités de transition entre les pairs sont égales, c'est le cas aussi pour les impairs entre-eux. Alors, entre les pairs et entre les impairs, ils ont une chance égale d'être atteint après un grand nombre d'itérations. Par conséquent, cela suggère que la distribution des probabilités à X200 converge vers une distribution uniforme distinguer selon si ils sont pairs ou impairs.

Loi uniforme	$\mathscr{U}$([a;b])	$P(X \in I) = \frac{long\ I}{b-a}$	$E(X) = \frac{a+b}{2}$	$V(X) = \frac{(b-a)^2}{12}$	Tous les intervalles de même longeur ont la même probabilité

On peut l'approximer à une loi normale

Loi Normale	$\mathcal{N}(\mu;\sigma)$	$\mathrm{P}(a \leq X \leq b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-½(\frac{x-\mu}{\sigma})^2}\,dx$	E(X)=μ	V(X)=σ^2	Nombreux facteurs rentrent en compte dont aucun n'est prépondérant. Approximation d'autres lois.

On peut, en réalité, en distinguer 2 avec des paramètres différents. Les états pairs et impairs.
Ils ont alors la même variance mais pas la même moyenne:

Moyenne pour les états pairs: $\mu = k/2$
Moyenne pour les états impairs: $\mu = (k/2) - 1$
Variance: $\sigma^2 = k^2/12$

Pour X200:
Moyenne pour les états pairs: $\mu = 100/2 = 50$
Moyenne pour les états impairs: $\mu = (100/2) - 1 = 49$
Variance: $\sigma^2 = k^2/12 = 100^2/12 = 833.33$