SCHOOL
OF
ENGINEERING
AND
MANAGEMENT

Télétrafique

Chaînes de Markov

Marcos Rubinstein

Index

1. Chaînes de Markov discrètes

1.1. Notions de base

Considérez un processus aléatoire discret dans le temps $\{X_m, m = 0, 1, 2, \ldots\}$, dans lequel la valeur de l'index m indique le moment auquel la variable aléatoire est observée.
En général, la valeur de la variable aléatoire au moment m (c'est à dire X_m) peut dépendre des valeurs antérieures $X_{m-1}, X_{m-2}, \ldots$.

Un exemple simple d'une variable aléatoire pourrait être deux états possibles de fonctionnement d'un appareil : “fonctionnement normal”, que nous pouvons appeler état 1, et “en panne”, que nous pouvons appeler état 0. Si nous utilisons m pour représenter le jour, nous pourrions observer par exemple la suite d'état suivante :
$\{X_0 = 1, X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 0, X_5 = 0\}$. Dans cette suite de moments d'observation, qu'on peut appeler une marche aléatoire, l'appareil est en l'état “fonctionnement normal” jusqu'au pas 3 et il passe à l'état “en panne” au pas 4.

Les chaînes de Markov sont des processus aléatoires spéciaux dans lesquels l'état X_m ne dépend que de l'état immédiatement antérieur X_{m-1} quelle que soit la valeur de m.

Les chaînes de Markov sont utilisées pour modéliser l'évolution des états dans des systèmes probabilistes dits markoviens. Dans l'exemple de l'appareil qui tombe en panne, il y avait seulement deux états : 0 et 1 pour “en panne” ou en “fonctionnement normal”. Dans le cas général, les systèmes peuvent prendre un nombre arbitrairement élevé d'états.

Considérez par exemple un système avec des états possibles $S = \{s_1, s_2, s_3, \ldots\}$. Les états sont typiquement choisis pour être soit $s_1 = 1, s_2 = 2, s_3 = 3, \ldots$ soit $s_1 = 0, s_2 = 1, s_3 = 2, \ldots$ en dépendant de ce qui convient le mieux dans un problème particulier. Si $X_m = i$, on dit que le système se trouve dans l'état i au pas d'observation m.
Supposons par exemple que nous modélisions une file d'attente au bureau de la poste. Le nombre de personnes dans la file est forcément un nombre entier non négatif. Si nous laissons X_n représenter le nombre de personnes dans la file à l'instant n, alors $X_n \in S = \{0, 1, 2, \ldots\}$.
L'ensemble S est appelé l'espace des états de la chaîne de Markov.

Voici une définition formelle des chaînes de Markov discrètes dans le temps :

Un processus aléatoire $\{X_m, m = 0, 1, 2, \ldots\}$ avec états possibles du system $S \subset \{0, 1, 2, \ldots\}$ est une chaîne de Markov si
$P(X_{m+1} = j \;\; X_m = i, X_{m-1} = i_{m-1}, \ldots, X_0 = i_0) = P(X_{m+1} = j \;\; X_m = i)$

Le nombre $P(X_{m+1} = j \;\; X_m = i)$ est appelé la probabilité de transition de l'état i à l'état j. On suppose que ces probabilités ne dépendent pas de m, c'est à dire qu'elles ne dépendent pas du temps. On peut donc écrire la probabilité p_{ij} de passer de l'état i à l'état j ainsi :
$p_{ij} = P(X_{m+1} = j \;\; X_m = i)$.

Puisque cette probabilité ne dépend pas du temps (autrement dit du moment d'observation), on peut écrire

$$
\begin{aligned}
p_{ij} &= P(X_1 = j \;\; X_0 = i) \\
&= P(X_2 = j \;\; X_1 = i) \\
&= P(X_3 = j \;\; X_2 = i) \\
&= \dots
\end{aligned}
$$

Répétons-le encore une dernière fois en d'autres mots : le processus fera une transition de i à j avec probabilité p_{ij} indépendamment de la valeur de m, qui représente le moment où cette transition se passe.

1.2. Diagramme de transition

Nous allons introduire une représentation graphique des chaînes de Markov, qui s'appelle **diagramme de transition**, en utilisant un example simple.

Une discothèque met chaque jour un de trois types de musique. Il y a des jours où elle met de la salsa, d'autres du rock&roll et encore d'autres jours où elle met du heavy metal. Les propriétaires de la discothèque ont leur propres règles pour décider quel type de musique ils vont faire écouter et danser à leurs clients chaque jour : le type de musique dépend seulement du type de musique de la soirée précédente. Par exemple, si aujourd'hui est un jour de heavy metal, la probabilité que la music de demain sera la salsa est 60%. On peut représenter ceci avec un diagramme en utilisant une flèche entre l'état présent et l'état futur comme illustré à la figure 1 ci-dessous :

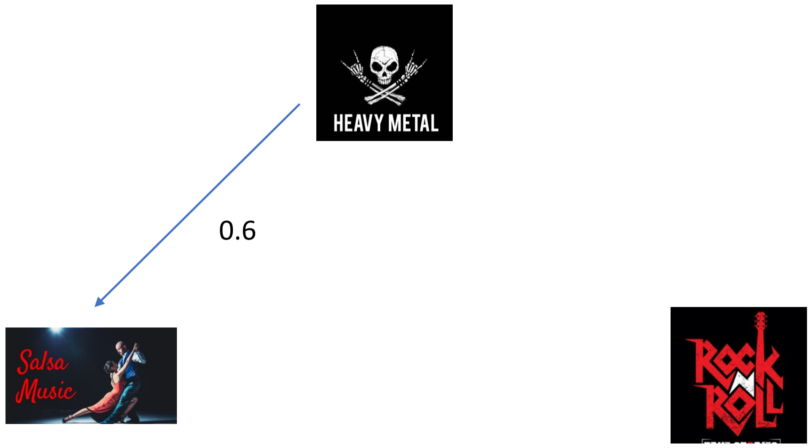


Figure 1. Discothèque avec un type de musique par jour et probabilité de transition entre heavy metal et salsa égal à 60%. Les autres probabilités ne sont pas représentées dans cette figure.

Il est aussi possible d'avoir le même état (le même type de musique) plusieurs jours d'affilé. Cette possibilité est représentée par une flèche en forme de boucle qui sort d'un état et qui retourne sur le même état. La figure 2 représente toutes les transitions dans notre exemple et les probabilités de transition y associées. Regardez le diagramme. Y a-t-il vraiment toutes le flèches ?

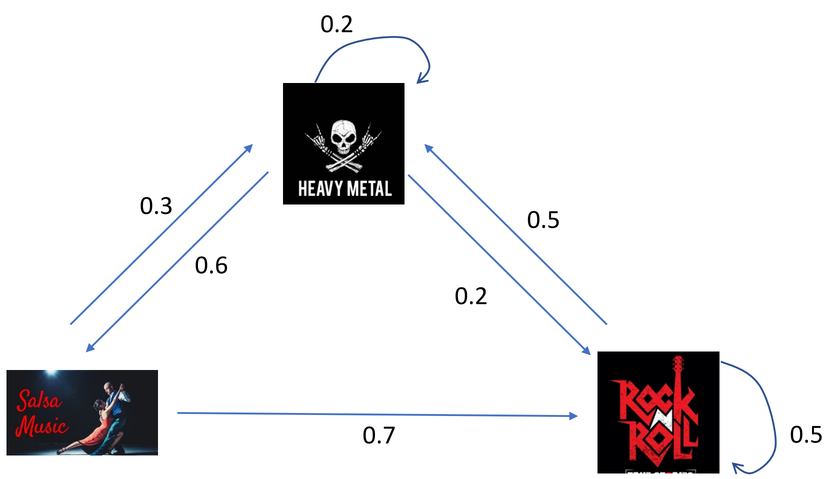


Figure 2. Discothèque avec un type de musique par jour avec les probabilités de transition

Les flèches qui manquent dans un diagramme de transition (c'est le nom que l'on donne à un tell diagramme dans le cadre des chaînes de Markov) sont celles qui correspondent à une probabilité nulle. Dans le cas de la figure 2, si la musique un jour donné est du Rock&Roll, la probabilité qu'elle soit de la salsa le lendemain est 0. La même chose est vraie pour l'autre flèche absente dans le diagramme de la figure 2 (pouvez-vous voir laquelle ?).

Comme vous voyez dans le diagramme, cette situation représente une chaîne de Markov puisque les probabilités de transition vers l'état d'observation suivant ne dépendent que de l'état actuel. Chaque état ne dépend donc de de l'état antérieur de la chaîne.

Supposons qu'un jour la discothèque met du Rock&Roll, que le jour d'après elle met du Heavy Metal et que le jour suivant elle met encore une fois du Rock&Rolle comme on voit à la figure 3.

Figure 3. Trois jours à la discothèque.

Quelle est la probabilité que la discothèque décide de mettre de la salsa le jour 4 ?
Essayez de donner la réponse avant de passer à la page suivante.

La réponse est 60% ou, exprimé autrement, 0.6. En effet, il suffit de voir à la figure 2 quelle est la probabilité de passer de Heavy Metal à salsa puisque, s'agissant d'un processus markovien, cette probabilité dépend seulement de l'état actuel et de l'état futur et pas de l'historique avant l'état actuel.

Une propriété importante à retenir pour ce type de chaîne de Markov est que la somme des poids des flèches qui sortent d'un état est égale à 1 puisqu'il s'agit de probabilités. Contrôlez que ceci est en effet le cas pour chacun des trois états à la figure 2.

Observons maintenant les types de musique dans notre chaîne de Markov pendant une période un peu plus longue en effectuant une marche aléatoire comme montré à la figure 4:

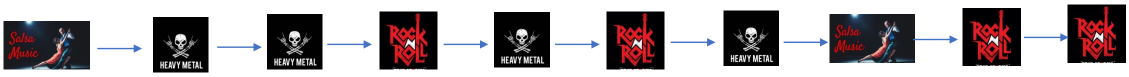
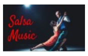


Figure 4. Dix jours de musique à la discothèque.

On voit qu'il a eu deux jours de salsa, quatre jours de Rock&Roll et encore quatre jours de Heavy Metal. Quelle est la probabilité de chaque état (de chaque type de musique) d'après ces 10 pas ?

On peut estimer ces probabilités en divisant le nombre de jours de chaque type de musique par le nombre total de jours dans cette marche aléatoire, qui est 10 :

$P($  $) = \frac{4}{10}$

$P($  $) = \frac{2}{10}$

$P($  $) = \frac{4}{10}$

Si nous répétons cet exercice avec le même nombre de pas, ou avec un nombre de pas différent (plus grand ou plus petit que 10), le résultat sera en général différent. Les questions qui se posent sont les suivantes : 1) est-ce que ces probabilités vont continuer à changer si nous observons de plus en plus de jours ? 2) sinon, vont-elles se stabiliser et converger à des valeurs fixes ?
3) si les probabilités de chaque état à long term se stabilisent, pouvons-nous en trouver les valeurs ?

Regardons dans les cas de la discothèque ce qui se passe à long terme en faisant une simulation :
Ecrivez un program en Python pour calculer la probabilité pour un grand nombre de pas. Pour faire ceci, vous devez pouvoir choisir le prochain type de musique à partir de la misique actuelle en utilisant les probabilités du diagramme de transition de la figure 2.

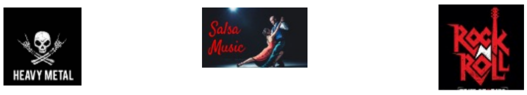
Quelles sont les valeurs des trois probabilités après 20 pas? Quelles sont ces valeurs après 50 pas ? Et après 1000 pas ? Trouvez les probabilités après 100'000 pas comme dernière exercice.

La distribution de probabilités que vous devriez avoir obtenue s'appelle la **distribution stationnaire**. Cette distribution ne change pas avec le temps pour cette chaîne de Markov.

1.3. Matrice de transition

Il y a une autre façon de trouver la distribution stationnaire en plus de celle qui implique une marche aléatoire par simulation d'un grand nombre de pas comme vous l'avez fait en programmant les transitions en Python à la fin de la section précédente.
Afin de présenter cette méthode, nous avons besoin d'une autre forme de représentation des chaînes de Markov : **la matrice de transition**.

La matrice de transition est construite avec les probabilités de transition. C'est une matrice carrée de dimension $M \times M$ où M est le nombre d'état dans la chaîne. La matrice de transition correspondant à l'exemple de la discothèque peut être construite en observant le diagramme de transitions de la figure 2 :



$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

Nous avons inclus des images à gauche et en haut de la matrice pour faciliter son interprétation mais ce type d'image n'est souvent pas inclus.

Les éléments de la matrice sont les probabilités de transition. Par exemple, la valeur du premier élément de la troisième ligne de la matrice (dont la valuer est 0.5) est la probabilité de transition de Rock&Roll à Heavy Metal (voir cette transition à la figure 2).

Les éléments dont la valeur est 0 correspondent aux transitions impossibles. Comment peut-on identifier ces transitions à la figure 2 (nous en avons déjà parlé mais, si vous ne vous souvenez pas, allez voir ces transitions à la figure) ?

Appelons P la matrice de transition. Cette matrice peut être écrite de manière générale comme suit

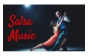
$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \tag{1}$$

Une propriété importante de cette matrice est que la somme des probabilités dans chacune des lignes est égale à 1. Réfléchissez à la raison de cette propriété en regardant à quoi correspondent les éléments d'une ligne dans le diagramme de transition de la figure 2.

Revenons encore une fois à l'exemple de la discothèque pour illustrer un autre concept. Supposons que nous savons que la musique sera la salsa aujourd'hui. Les probabilités des trois types de musique pour aujourd'hui sont donc

$P($ 

$) = 0$

$P($  $) = 1$

$P($  $) = 0$

Ecrivons maintenant ces probabilités sous la forme d'un vecteur que nous appellerons $\pi^{(0)}$:

$$\pi^{(0)} = [0 \;\; 1 \;\; 0] \tag{1}$$

Le superscript "(0)" indique qu'il s'agit de la condition à l'état 0, donc à l'état initiale.
Le premier élément de ce vecteur $\pi^{(0)}$ indique que la probabilité d'entendre du Heavy Metal à la discothèque aujourd'hui est 0. Le deuxième élément indique pour sa part que la probabilité d'entendre de la salsa est 1 et le troisième élément que la probabilité de Rock&Roll est 0.

Les probabilités de chacun des types de musique un jour plus tard peuvent être calculées en multipliant le vecteur $\pi^{(0)}$ par la matrice P :

$$\pi^{(1)} = \pi^{(0)} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix} \tag{2}$$

Remplaçant $\pi^{(0)}$, en obtient

$$\pi^{(1)} = [0 \;\; 1 \;\; 0] \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix} = [0.3 \;\; 0 \;\; 0.7] \tag{3}$$

Observant cette équation, nous pouvons affirmer qu'après un jour, nous avons 30% de probabilité de Heavy Metal, 0% de probabilité de salsa et 70% de probabilité de Rock&Roll. Ce n'est pas étonnant que la probabilité d'avoir de la salsa demain soit zéro puisque selon le diagramme de transitions, ce n'est pas possible de mettre de la salsa deux jours de suite.

Nous pouvons maintenant trouver la probabilité de chacun des types de musique auprès-demain en multipliant par P le vecteur de probabilités de demain $\pi^{(1)}$ que nous venons de calculer :

$$\pi^{(2)} = \pi^{(1)} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix} \tag{4}$$

Remplaçant le vecteur $\pi^{(1)}$ de l'équation (3), en obtient

$$\pi^{(2)} = [0.3 \;\; 0 \;\; 0.7] \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix} = [0.41 \;\; 0.18 \;\; 0.41] \tag{5}$$

Si dans l'équation (4) on remplace l'équation (3), nous obtenons

$$\pi^{(2)} = \pi^{(0)} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix}^2 \qquad (6)$$

Oú P est élevé à la deuxième puissance.

Nous pouvons utiliser ce résultat pour généraliser le calcul des probabilités qu'une chaîne de Markov se trouve dans chacun de ces états à l'état n :

$\pi^{(n)} = \pi^{(0)} P^n$

où P^n est la matrice P élevée à la puissance n.

De manière plus générale, la probabilité qu'une chaîne de Markov passe de l'état i à l'état j en exactement m pas est donnée par les éléments de la matrice

$(P_{ij}^m) = P^m$

Avant de présenter une autre façon d'utiliser ce que nous venons de voir pour trouver la distribution stationnaire de probabilités, voici quelques exercices.

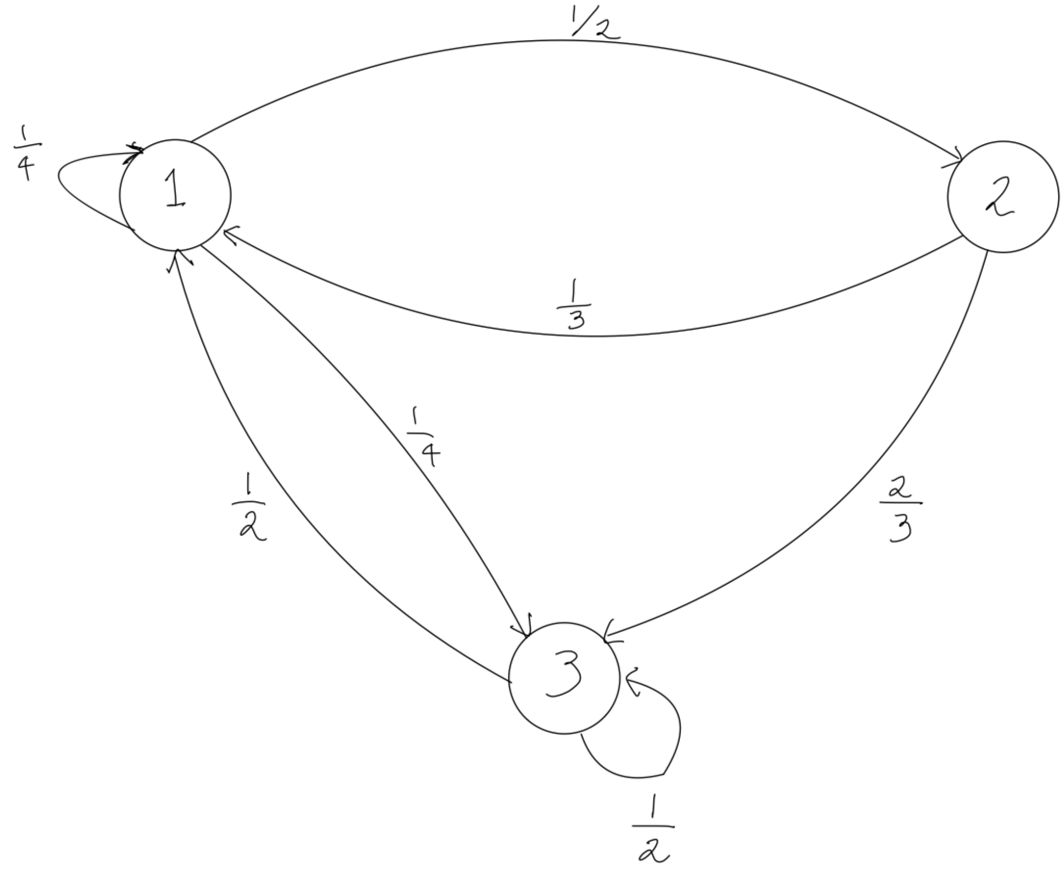
Exercices 1

1) Utilisez Python pour calculer la distribution des probabilités des trois musiques après 10 pas, puis 20 fois et ensuite 30 fois si la distribution de probabilités initiale est donnée par le vecteur $\pi^{(0)}$ donnée précédemment.

2) Trouvez les valeurs de probabilités de transition marquées “?” dans la matrice de transition suivante et dessinez le diagramme de transition correspondant :

$$P = \begin{bmatrix} ? & \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{2} & ? & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & ? & 0 \\ ? & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

3) Considérez la chaîne de Markov suivante :



a) Trouvez $P(X_4 = 3 \ \ X_3 = 2)$.

b) Trouvez $P(X_3 = 1 \quad X_2 = 1)$.

c) Trouvez $P(X_0 = 1, X_1 = 2)$ sachant que $P(X_0 = 1) = \frac{1}{3}$.

d) Trouvez $P(X_0 = 1, X_1 = 2, X_2 = 3)$ sachant que $P(X_0 = 1) = \frac{1}{3}$.

1.4. Calcul de la distribution stationnaire

La distribution de probabilités $\pi^{(n+1)}$ au pas $n+1$ peut être calculée à partir de la distribution $\pi^{(n)}$ au pas n en utilisant l'équation suivante :

$$\pi^{(n+1)} = \pi^{(n)}P$$

Après un grand nombre de pas, si la distribution de probabilités devient stationnaire (ce n'est pas le cas pour toutes les chaînes de Markov comme nous verrons plus bas), elle convergera vers une distribution que nous appellerons π. Nous pouvons donc écrire

$$\pi = \pi P \qquad (1)$$

Notez que cette équation ne dit pas que les états deviennent constant. C'est seulement leur probabilités qui arrivent à un point d'équilibre et elle ne changent plus. Le vecteur π dans l'équation (1) est ce que l'on appelle un eigenvector associé à la transformation produite par la matrice P.

Essayons d'utiliser l'équation (1) pour trouver la distribution de probabilités π avec un example :

Supposons que la matrice P est la suivante :

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

- Dessinez d'abord le diagramme de transition correspondant à cette matrice de transition.

- Essayez maintenant de résoudre le système d'équations posé par $\pi = \pi P$ pour trouver les valeurs de x et de y qui sont les probabilités stationnaires dans $\pi = [x \quad y]$.

  Souvenez-vous que les sommes des probabilités sortant d'un état doit être égale à 1. Ceci revient à dire que

  $$p_{11} + p_{12} = 1$$

  et

  $$p_{21} + p_{22} = 1$$

Si vous n'arrivez pas à trouver la solution avec le système d'équations $\pi = \pi P$, l'essai n'a pas été en vain. En effet, si vous remarquez que les deux équations que vous avez trouvées sont linéairement dépendantes, il suffit de résoudre le système d'équations représenté par une d'entre les deux équations initiales et une nouvelle équation se basant sur le fait que la somme des probabilités stationnaires doit elle aussi être égale à 1:

$\pi[1] + \pi[2] = 1$. L'équation est donc $x + y = 1$

Les chaînes de Markov peuvent avoir plus d'une distribution stationnaire. Elles peuvent aussi n'en avoir aucune.
Regardons les conditions qui font qu'une chaîne avec matrice de transition P ait une distribution stationnaire π. Les trois conditions suivantes doivent être satisfaites :

- $\pi \cdot P = \pi$
- La somme des probabilités dans le vecteur π est égale à 1.
- Toutes les probabilités dans π sont plus grandes ou égales à 0.

Voyons quelques exemples.

Exemple 1

Supposons que les suisses et les suissesses qui font leur service militaire font des exercices de tir avec un des armes de dernière génération héritées de Guillaume Tell et qu'elles ont le choix entre deux cibles comme montré à la figure ci-dessous.

On observe pendant longtemps les tirs de la recrute de la figure et on se rend compte que, tout comme son arrière grand-père (justement un tell Guillaume Tell), elle ne rate jamais les cibles. On observant quantitativement ses prouesses, on conclut que si elle vise cette fois-ci la cible 1, la probabilité qu'elle vise la même cible la prochaine fois est 80%. D'autre part, lorsqu'elle vise la cible 2, la probabilité qu'elle vise la cible 1 la prochaine fois est 60%.

Puisque ces probabilités ne dépendent que de l'état actuel, il s'agit d'un processus markovien. Dessinez le diagramme de transitions :

Ecrivez la matrice de transitions :

$$P = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

Utilisez la méthode des eigenvalues pour trouver les valeurs des probabilités stationnaires :

$$\pi \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} = \pi$$

Avec $\pi = [x \quad y]$, écrivez deux équations de la forme suivante avec x et y comme inconnues :

$$ax + by = x$$
$$cx + dy = y$$

Ecrivez ces équations ici :

$\_\_\_x + \_\_\_y = x$

$\_\_\_x + \_\_\_y = y$

Comme nous avons vu précédemment, il faut aussi utiliser la relation entre les probabilités x et y :

$x + y = 1$

Résolvez ces équations simultanées et notez les résultats :

$x =$

$y =$

Vous pouvez maintenant écrire le vecteur de probabilités stationnaires π utilisant ces deux valeurs :

$\pi = [\qquad\qquad\qquad\qquad\qquad\qquad]$

Pour ce type de chaîne de Markov (nous verrons les conditions nécessaires plus tard), il est possible de trouver cette distribution stationnaire en utilisant une autre méthode qui se base sur le fait que les probabilités convergent vers la distribution stationnaire après un grand nombre de pas. Vous pouvez initialiser la chaîne avec n'importe quelle distribution de probabilités (pour autant que la somme soit égale à 1).

Initions le calcule avec $\pi^0 = [1 \quad 0]$

Calculez maintenant utilisant python, Matlab ou autre logiciel de calcul $\pi \cdot P^n$ pour une grande valeur de n, par exemple 1000 (c'est souvent possible d'arriver à la distribution stationnaire avec un exposant plus petit):

$$[1 \quad 0]\begin{bmatrix} \qquad\qquad & \qquad\qquad \\ \qquad\qquad & \qquad\qquad \end{bmatrix}^{1000} = [\qquad\qquad\qquad\qquad\qquad\qquad]$$

Vous devriez avoir obtenu le même résultat qu'avec la méthode précédente.

Example 2

Ce deuxième exemple a pour but de montrer que même si une distribution stationnaire existe, on ne peut pas forcément y arriver avec une marche aléatoire avec un grand nombre de pas.
Nous nous baserons sur l'exemple précédent mais, cette fois-ci, on donne l'ordre à Madame Tell d'alterner entre les deux cibles d'une fois à la suivante.

Dessinez le nouveau diagramme de transitions et écrivez la nouvelle matrice de transition :

$$P = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

Appliquez à cette nouvelle chaîne la méthode des eigenvalues pour trouver les valeurs des probabilités stationnaires :

$$\pi \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} = \pi$$

Ecrivez encore les deux équations qui en résultent utilisant $\pi = [x \quad y]$:

$\_\_\_x + \_\_\_y = x$

$\_\_\_x + \_\_\_y = y$

La relation entre les probabilités x et y est encore :

$x + y = 1$

Obtenez les valeurs de x et de y,

$x =$

$y =$

et écrivez le vecteur de probabilités stationnaires :

$\pi = [\qquad\qquad\qquad\qquad\qquad\qquad]$

Comme dans l'exemple 1, essayons de calculer la distribution stationnaire avec la méthode de la puissance de la matrice de transition :

Initions la marche encore une fois avec $\pi^0 = [1 \quad 0]$

et calculons encore

$\pi \cdot P^n$

pour une grande valeur de n. Utilisez d'abord 1000 :

$$[1 \quad 0]\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{1000} = [1 \quad 0] \neq [\frac{1}{2} \quad \frac{1}{2}]$$

Le résultat dans ce cas-ci est différent que celui obtenu avec la méthode des eigenvecteurs!

Attention: le vecteur $[\frac{1}{2} \quad \frac{1}{2}]$ est le résultat de la méthode des eigenvecteurs que vous avez calculé tout à l'heure.

En plus, si nous calculons la distribution avec un exposant impair, nous obtenons une réponse différente :

$$[1 \quad 0]\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{1001} = [0 \quad 1]$$

On peut encore faire un autre test avec cette matrice de transitions. Cette chaîne de Markov ne converge définitivement à l'état stationnaire $[\frac{1}{2} \quad \frac{1}{2}]$ que si les probabilités initiales sont $[\frac{1}{2} \quad \frac{1}{2}]$. En effet,

$$[\frac{1}{2} \quad \frac{1}{2}]\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{n} = [\frac{1}{2} \quad \frac{1}{2}]$$

pour n'importe quelle valeur de n

Voyons maintenant comment déterminer si une chaîne de Markov a une distribution stationnaire unique et que la chaîne converge à cette distribution indépendamment des conditions initiales lorsque l'on fait une marche avec un grand nombre de pas qui tendent vers l'infini.

Les conditions sont les suivantes (pour les définitions, référez-vous é la section 1.9 :

Si une chaîne de Markov est **irréductible** et **apériodique**, et s'il existe une distribution stationnaire π, alors la chaîne converge vers cette distribution lorsque le nombre de pas tend vers l'infini, quels que soient les états de départ.

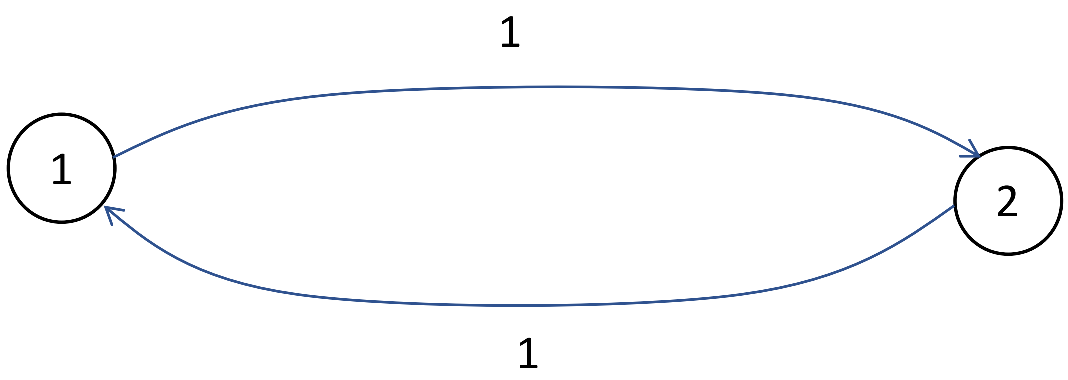
Voici la définition des chaînes de Markov irréductibles :

Si tous les états d'une chaîne de Markov sont accessibles depuis les autres états, on dit qu'ils appartiennent à une même ***classe*** *et la chaîne est dite* ***irréductible****.*

Voyons maintenant ce qu'une chaîne de Markov apériodique.

1.5. Périodicité

La chaîne que nous avons utilisée à l'exercice 2 ci-dessus recopiée ci-dessous pour convenance, se comporte de manière périodique.



***En effet, si l'on commence à ‘état 1, la chaîne passe au premier pas forcément à l'état 2, puis, au prochain pas, de nouveau à l'état 1. La chaîne est à l'état $X_m = 1$ pour toute valeur pair de m et à l'état 2 pour toute valeur impaire de m. Cela a comme conséquence que lorsque $m \to \infty$, la probabilité de l'état 1 est égale à 1 pour les valeurs pairs et cette probabilité est égale à zéro pour les valeurs impaires de m. La chaine ne converge donc pas.

1.6. La période d'un état

Intuitivement, la période est définie de sorte que le temps nécessaire pour passer de l'état i de nouveau à l'état i soit toujours un multiple de cette période.
Dans l'exemple ci-dessus, la chaîne peut revenir à l'état 1 après 2 pas, 4 pas, 6 pas, 8 pas, . . . La période de l'état 1 est donc 2.
En général, la chaîne peut partir depuis l'état i et y retourner en t pas si
$(P^t)_{ii} > 0$, c'est à dire, si la probabilité de retourner à l'état i ayant parti du même état en t pas n'est pas 0. Ceci amène la définition suivante :

La période $d(i)$ de l'état i est

$$d(i) = gcd\{t\ (P^t)_{ii} > 0\}$$

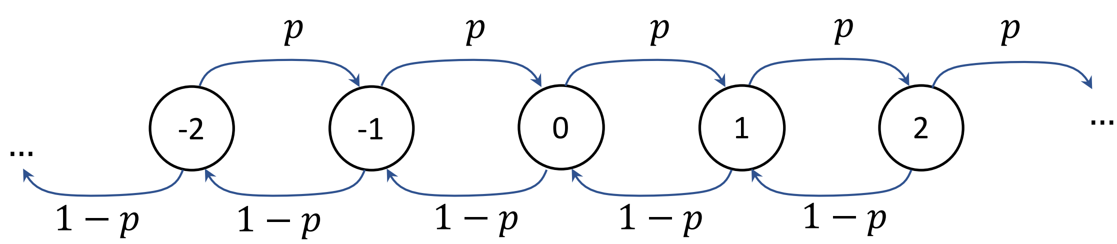
c'est-à-dire, le plus grand diviseur commun des temps pour lesquels le retour est possible.

L'état est dit **périodique** si $d(i) > 1$.
Attention : si la période est 1, l'état n'est pas périodique! En effet, si $d(i) = 1$, l'état i est appelé **apériodique**.

On va seulement s'intéresser aux états apériodiques pour la convergence à un état stationnaire.

Exemples

1. Considérez la chaîne de Markov suivante :



a) Quelle est sa période par rapport à l'état 0 ?

Les nombres de pas pour lesquels la probabilité de retourner à l'état 0 n'est pas zéro sont :
- 2 lorsqu'on passe à l'état 1 à droite et on retourne à l'état 0, ou vers la gauche, lorsque la chaîne passe à l'état -1 et retourne à l'état 0.
- 4 quand la chaîne suit les états suivants : $0 \to 1 \to 2 \to 1 \to 0$ ou vers la gauche $0 \to -1 \to -2 \to -1 \to 0$

et de la même façon pour le nombre pairs plus grands que 4. Nous pouvons donc écrire la période $d(0)$ pour l'état 0 comme suit

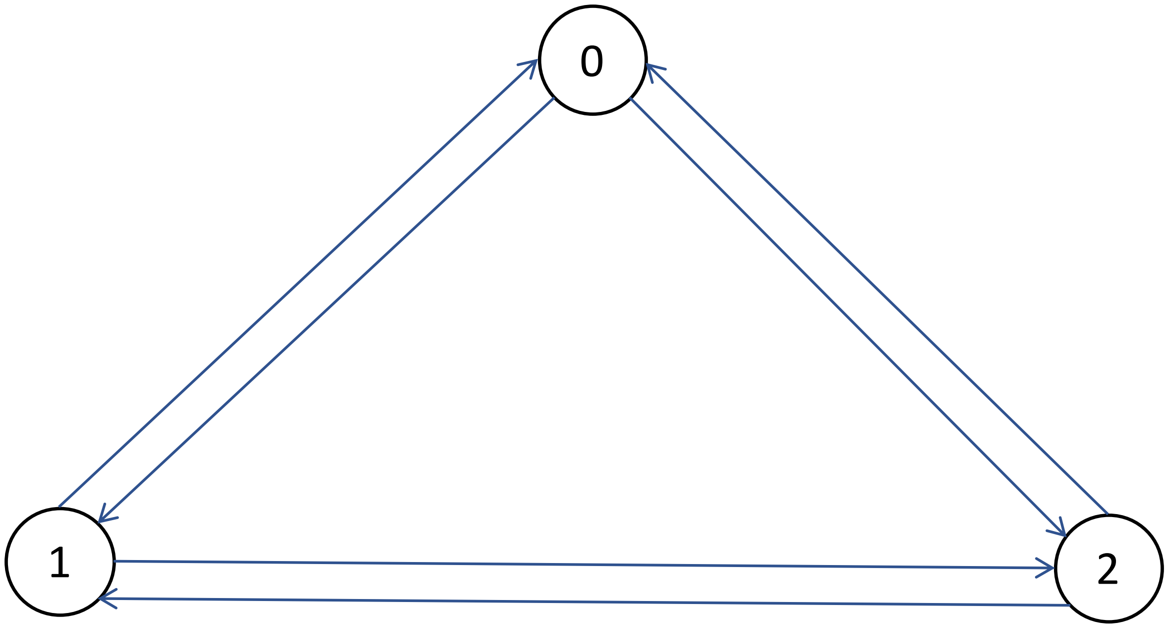
$$d(0) = gcd\{2,4,6,\ldots\} = 2$$

b) Est-ce que la chaîne est irréductible ?
Oui puisque tous les états communiquent entre eux.

c) La chaîne, est-elle périodique ou apériodique ?
Puisque la chaîne est irréductible et que la période d'un des états est 2, la chaîne est périodique.

2. Considérez la chaîne de Markov suivante :

a) Quelle est sa période par rapport à l'état 0 ?



Les nombres de pas pour lesquels la probabilité de retourner à l'état 0 n'est pas zéro sont :

- 2 pour les transitions $0 \to 1 \to 0$ et aussi $0 \to 2 \to 0$
- 3 quand la chaîne suit les états suivants : $0 \to 1 \to 2 \to 0$ ou bien $0 \to 2 \to 1 \to 0$
- 4 quand la chaîne suit les états suivants : $0 \to 1 \to 2 \to 1 \to 0$ ou bien $0 \to 2 \to 1 \to 2 \to 0$
- 5 quand la chaîne suit les états suivants : $0 \to 1 \to 2 \to 1 \to 2 \to 0$ ou $0 \to 2 \to 1 \to 2 \to 1 \to 0$

De façon similaire, on peut aussi trouver la manière de retourner à l'état 1 en $6, 7, 8, \ldots$ pas. Nous pouvons donc écrire la période $d(0)$ pour l'état 0 comme suit

$$d(0) = gcd\{2,3,4,5,6,7,8...\} = 1$$

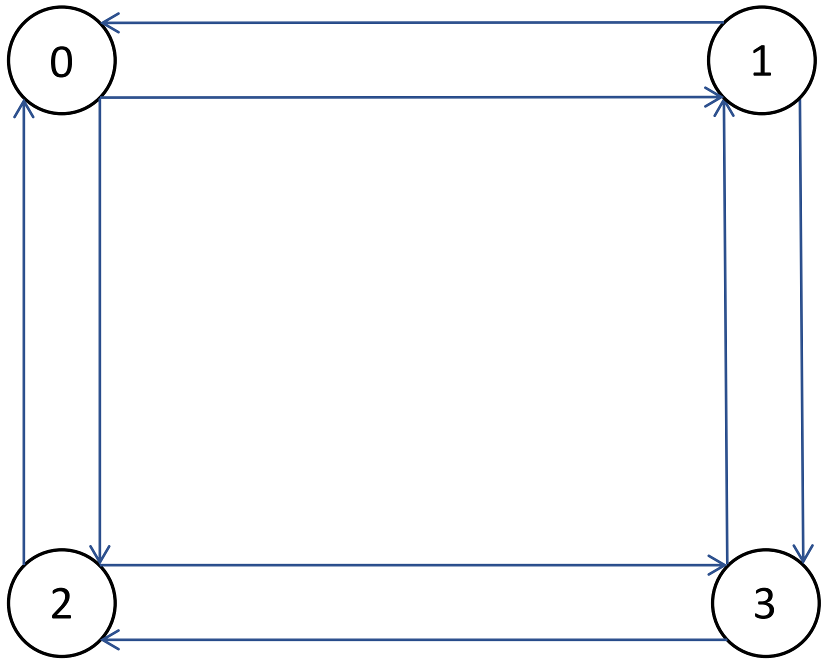
b) Est-ce que la chaîne est irréductible ?

Oui puisque tous les états communiquent entre eux.

c) La chaîne, est-elle périodique ou apériodique ?

Puisque la chaîne est irréductible et que la période d'un des états est 1, la chaîne est apériodique.

3. Considérez la chaîne de Markov suivante :



a) Quelle est sa période par rapport à l'état 0 ?

***Nous pouvons trouver les façons de retourner à l'état 0 en $2, 4, 6, \ldots$ pas. La période $d(0)$ pour l'état 0 est donc

$$d(0) = gcd\{2{,}4{,}6{,}8...\} = 2$$

Un commentaire qui est valable pour toutes les chaînes de Markov. Si on peut retourner à un état en un nombre de pas n, on peut aussi y retourner en $2n$, $3n$, $4n$ et ainsi de suite puisqu'on peut inclure les cas où on retourne à l'état plusieurs fois. Pour clarifier un peu ce commentaire, voici une des façons de retourner à l'état 0 en 8 pas :

$$0 \to 1 \to 0 \to 1 \to 0 \to 1 \to 0 \to 1 \to 0$$

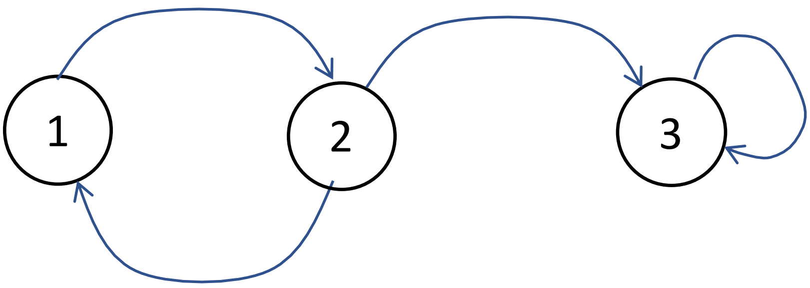
b) Est-ce que la chaîne est irréductible ?

Oui puisque tous les états communiquent entre eux.

c) La chaîne, est-elle périodique ou apériodique ?

Puisque la chaîne est irréductible et que la période d'un des états est 2, la chaîne est périodique.

4. Considérez la chaîne de Markov suivante :



a) Quelle est sa période par rapport à l'état 1 ?

Nous pouvons trouver des façons de retourner à l'état 1 en $2, 4, 6, \ldots$ pas. La période $d(0)$ pour l'état 0 est donc

$$d(0) = gcd\{2,4,6,8...\} = 2$$

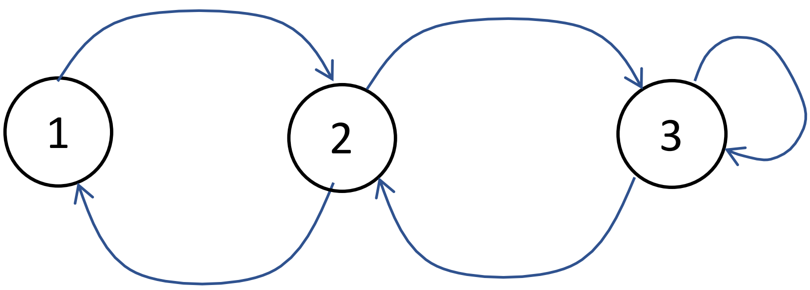
b) Est-ce que la chaîne est irréductible ?

Non, puisque pas tous les états communiquent entre eux. La chaîne est donc réductible.

c) La chaîne, est-elle périodique ou apériodique ?

Puisque la chaîne est réductible, nous ne pouvons pas parler de périodicité pour la chaîne entière. Cependant, pour la classe composée des états 1 et 2, la période étant 2, cette classe est périodique.

5. Considérez la chaîne de Markov suivante :



a) Quelle est sa période par rapport à l'état 1 ?

Nous pouvons trouver les façons de retourner à l'état 1 en $2, 4, 5, 6, 7\ldots$ pas. La période $d(0)$ pour l'état 0 est donc

$$d(0) = gcd\{2,4,5,6,7,8...\} = 1$$

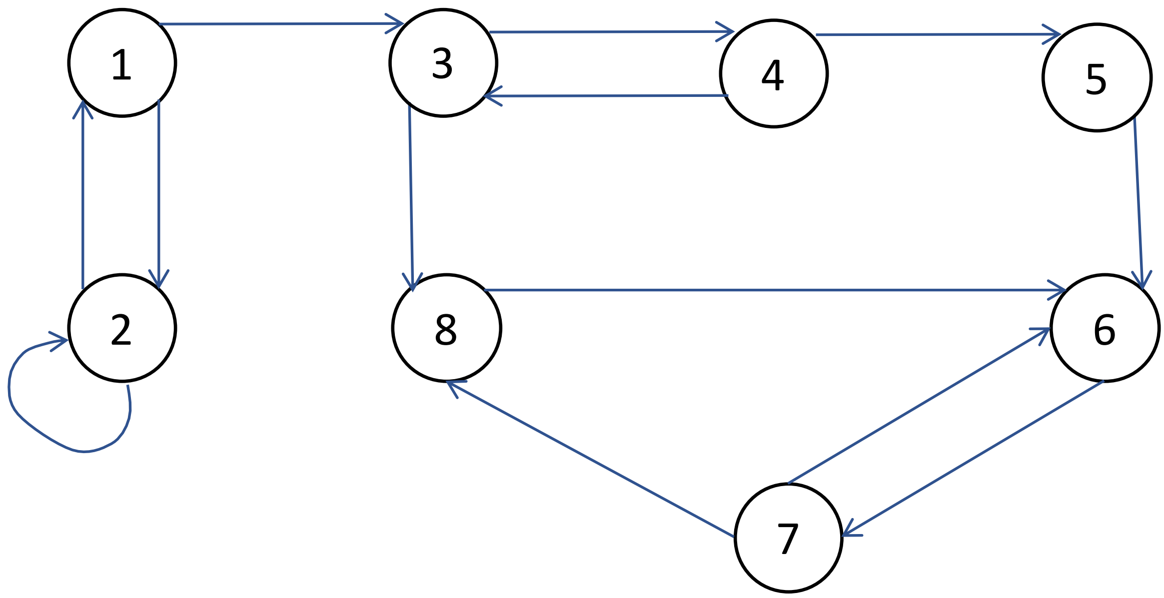
b) Est-ce que la chaîne est irréductible ?

Oui, puisque tous les états communiquent entre eux.

c) La chaîne, est-elle périodique ou apériodique ?

Puisque la chaîne est irréductible et que la période d'un des états est 1, la chaîne est apériodique.

6. Considérez la chaîne de Markov suivante :



Dans cet exercice, on voit que la chaîne n'est pas irréductible puisqu'il y a des états qui ne communiquent avec les autres (par exemple l'état 5).

a) Combien de classes y a-t-il dans cette chaîne ?

3

b) Donnez chacune des classes

$Class\ 1 = \{Etat\ 1, Etat\ 2\}$
$Class\ 2 = \{Etat\ 3, Etat\ 4\}$
$Class\ 3 = \{Etat\ 6, Etat\ 7, Etat\ 8\}$

c) Quelle est la période par rapport à l'état 1 ?

$d(1) = gcd\{2,3,4,5,6,7,8...\} = 1$

d) Quelle est la période par rapport à l'état 3 ?

$d(3) = gcd\{2,4,6,8...\} = 2$

e) Quelle est la période par rapport à l'état 6 ?

$d(0) = gcd\{2,3,4,5,6,7,8...\} = 1$

f) La $Class\ 1$, est-elle périodique ou apériodique ?

Apériodique, puisque la classe est irréductible et la période est 1.

g) La $Class\ 2$, est-elle périodique ou apériodique ?
Périodique, puisque la période est 2

h) La $Class\ 3$, est-elle périodique ou apériodique ?

Apériodique, puisque la classe est irréductible et la période est 1.

1.7. Critères pour trouver la périodicité d'une chaîne de Markov irréductible ou d'une classe dans une chaîne réductible

Une chaîne est apériodique si une des ces trois conditions est remplie :

1. Si la probabilité de rester à un des états de la chaîne est plus grande que zéro
2. Si on peut commencer à l'état i et y retourner en un nombre m ou en un nombre l de pas, et si m et l sont co-primes (c'est à dire que le seule facteur commun est 1).
3. S'il existe un nombre entier n positif tell que les éléments de la matrice de transition P^n sont tous strictement positifs.

1.8. Convergence

Si deux états i et j communiquent entre eux ($i \leftrightarrow j$), les deux états ont la même période : $d(i) = d(j)$.

De ce fait, nous pouvons conclure que si une chaîne de Markov est irréductible et qu'un de ces état et apériodique, tous les états de la chaîne sont apériodiques.

Une chaîne de Markov de ce type est appelée une chaîne irréductible et apériodique.

Exercices 2

1) Calculez le probabilités stationnaires pour l'example de la discothèque que nous avons utilisé au début de ce support. Pour rappel, la matrice de transition est

$$P = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0 & 0.7 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

Aide pour la solution de cet exercice :

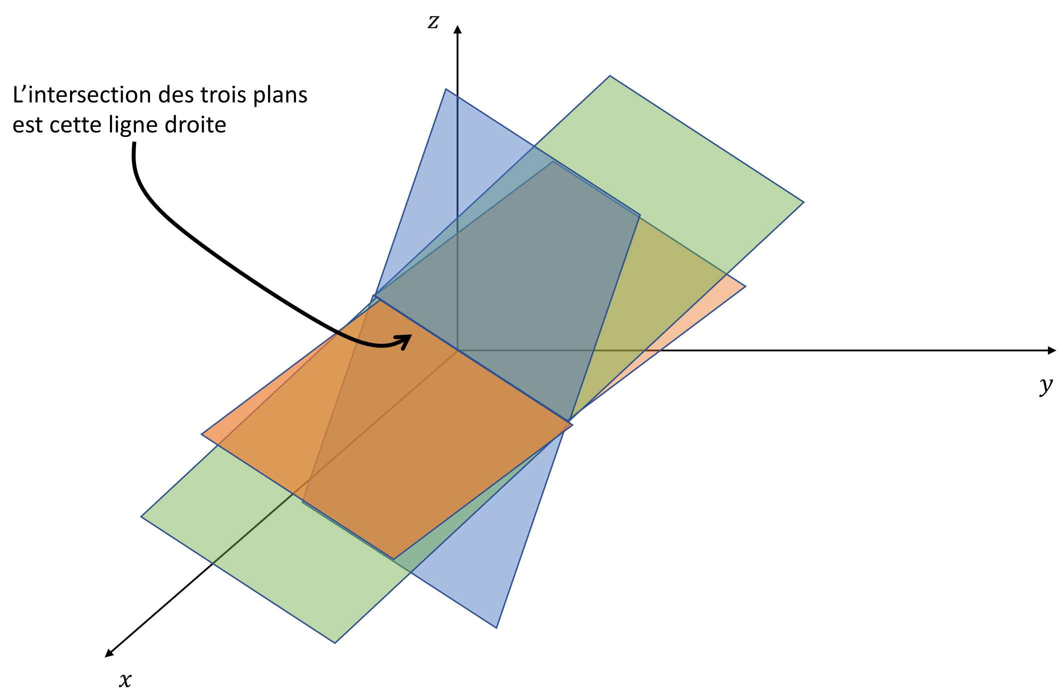
La solution de cet exercice implique la résolution d'un système de 4 équations avec tours inconnues. Dans le cas de chaînes de Markov avec une distribution stationnaire, les trois premières équations sont de la forme suivante :

$$ax + bx + cz = 0$$
$$dx + ex + fz = 0$$
$$gx + hx + pz = 0$$

où les coefficients a, b, c, d, e, f, g, h et p sont des éléments de la matrice de transitions et x, y et z sont les inconnues (les probabilités stationnaires).

Chacune de ces équations représente un plan dans l'espace. Puisque les équations n'ont pas de terme indépendant, tous les plans passent par l'origine du système de coordonnées. On peut voir que les plans contiennent effectivement l'origine en remplaçant les coordonnées e l'origine dans n'importe laquelle des équations ($x = 0$, $y = 0$ et $z = 0$).

La situation est illustrée à la figure ci-dessous, dans laquelle les plans ont été représentés avec des couleurs différentes.



Puisque l'intersection est une ligne (en l'occurrence droite), il y a un nombre infini de solution des trois premières équations, les solutions étant tous les points dans la droite. La quatrième équation, celle qui simule que la somme des probabilités de chacun de trois états doit être égale à 1,

$$x + x + z = 1$$

va nous permettre de déterminer lequel des points de la droite d'intersection est la solution des quatre équations. On peut procéder comme suit :

2) Trouver l'équation paramétrique de la droite.

Puisque la droite est l'intersection de n'importe quelle pair parmi les trois plans, prenons les deux premiers, dont les équations sont répétées ici par convenience :

$$ax + bx + cz = 0$$

$$dx + ex + fz = 0$$

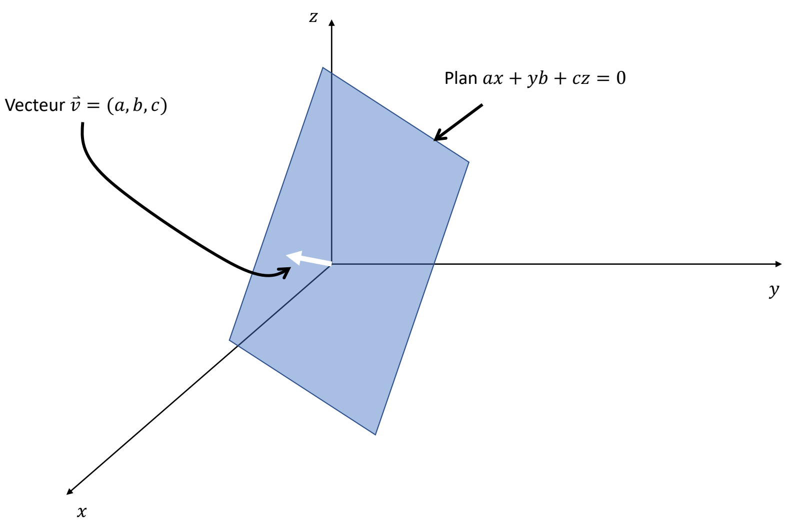
La forme de l'équation paramétrique d'une ligne droite qui passe par l'origine est

$$(x, y, z) = t \cdot (v_x, v_y, v_z)$$

où un point (x, y, z) dans la droite peut être trouvé en multipliant un paramètre t par un vecteur $\overrightarrow{v_d} = (v_x, v_y, v_z)$ dans la direction de la droite.

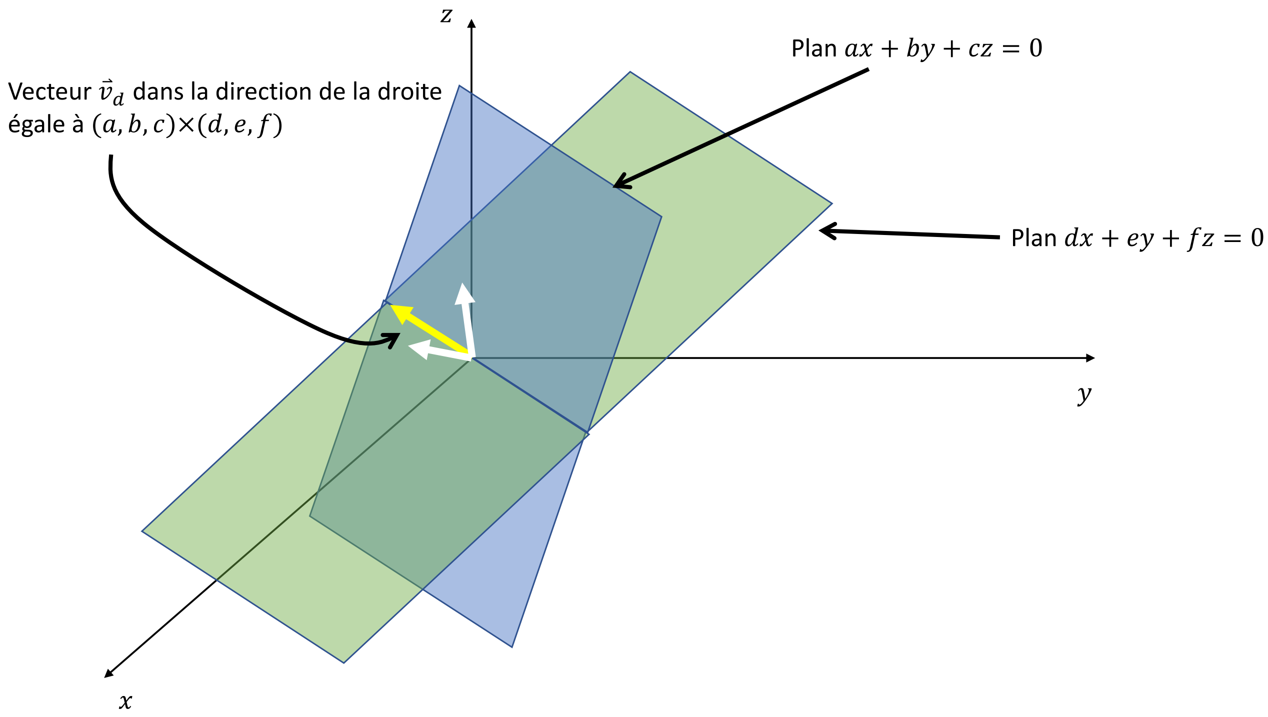
Mais, comment trouver le vecteur dans la direction de la droite ?

Nous allons le faire en nous basant sur le fait que, dans l'équation d'une droite, les coefficients de x, y et z sont le composants d'un vecteur perpendiculaire au plan de l'équation. Ainsi, pour le premier plan, le vecteur perpendiculaire est (a, b, c) comme illustré à la figure ci-dessous,



De la même façon, les coefficients de l'équation du deuxième plan forment un vecteur perpendiculaire au deuxième plan.

On peut voir à la figure ci-dessous que le vecteur $\overrightarrow{v_d}$ est perpendiculaires à ces deux vecteurs et il peut donc être calculé comme le produit vectoriel $(a\ , b\ , c) \times (d\ , e\ , f\ )$ des deux vecteurs.



Nous pouvons maintenant écrire l'équation paramétrique de la droit :

$(x, y, z) = t \cdot (a\ , b\ , c) \times (d\ , e\ , f\ )$

Il est maintenant simple de trouver la valeur de t pour que $x + x + z = 1$ soit satisfait. Cette valeur de t peut ensuite être utilisée dans $(x, y, z) = t \cdot (a\ , b\ , c) \times (d\ , e\ , f\ )$ pour trouver les valeurs de x, y et z.

3) Comparez le résultat avec les résultat obtenus à l'exercice 1. de la série d'exercices 1.

1.9. Quelques définitions

L'état j est **accessible** depuis l'état i si et seulement si il existe un entier $n \geq 0$ pour lequel

$p_{ij}^{n} > 0.$

Dit autrement, l'état j est **accessible** depuis l'état i si c'est possible d'arriver à l'état j depuis l'état i en un nombre fini de pas. On peut écrire cette relation de manière symbolique comme ainsi: $i \rightarrow j$

S deux états sont accessibles mutuellement, on dit qu'ils **communiquent**. On peut écrire cette relation $i \longleftrightarrow j$.

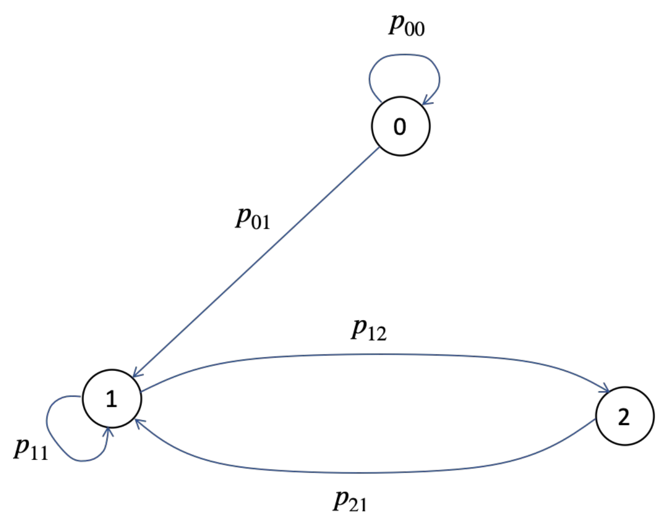
Si tous les états d'une chaîne de Markov sont accessibles depuis les autres états, on dit qu'ils appartiennent à une même **classe** et la chaîne est dite **irréductible**. Une **classe de communication** (communication class en anglais) est donc un ensemble d'états qui communiquent entre eux dans une chaîne de Markov.

Un état est dit **récurrent** si la probabilité d'y retourner est égale à 1.

Un état est dit **transitoire** si la probabilité d'y retourner est plus petite que 1.

Un état est dit **absorbant** si la probabilité de rester dans cet état une fois y arrivé est égale à 1.

Considérons la chaîne de Markov suivante :



Dans ce diagramme, lorsqu'une probabilité de transition est zéro, comme par exemple de l'état 1 à l'état zero ou entre les états 0 et 2, la flèche n'est pas dessinée.

Essayez de répondre aux questions suivantes :

1) L'état 0, est-il **accessible** depuis l'état 1 ?

2) L'état 0, est-il **accessible** depuis l'état 2 ?

3) L'état 1, est-il **accessible** depuis l'état 2 ?

4) L'état 1, est-il **accessible** depuis l'état 0 ?

5) Les états 1 et 2, **communiquent**-ils ?

6) Les états 1 et 0, **communiquent**-ils ?

7) Les états 0 et 2, **communiquent**-ils ?

8) Cette chaîne de Markov, est-elle **irréductible** ?

9) Lesquels des états de cette chaîne de Markov sont **récurrents** ?

10) Lesquels des états de cette chaîne de Markov sont **transitoires** ?

11) Lesquels des états de cette chaîne de Markov sont **absorbants** ?

12) Les états 0, 1 et 2, appartiennent-ils à une **classe** de communication ?

Faisons maintenant une marche aléatoire dans ce diagramme commençant à l'état 0 :
$0 \to 0 \to 0 \to 1 \to 1 \to 2 \to 1 \to 2 \ldots$

Analysons cette marche aléatoire. Les deux premières transitions, nous restons à l'état zéro. Nous passons ensuite à l'état 1. Ce pas était inévitable puisque la probabilité de transition de l'état 0 à l'état 1 n'est pas zéro ($p_{01} \neq 0$). Une fois que la chaîne sort de l'état 0, il n'y a aucune manière d'y retourner comme on peut voir dans le diagramme de transition. Cet état a une probabilité de ne pas pouvoir y retourner qui est plus petite que 1. Il est donc un état transitoire.

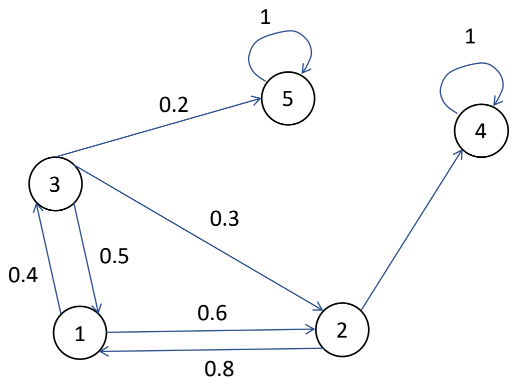
D'un autre côté, si l'on regarde l'état 1, on voit que si l'on en sort lors d'une transition vers l'état 2, il est sûr à 100% que l'on reviendra à l'état 1 dans le future. Etant donc un état dont la probabilité de revenir est égale à 1, il s'agit d'un état **récurrent.** On peut voir que l'état 2 est aussi un état récurrent.

1.10. Quelques observations

Un état **récurrent** est visité un nombre infini de fois.
Un état **transitoire** n'est visité qu'un nombre fini de fois.
Si on démarre la chaîne de Markov dans un état **récurrent**, on y reviendra un nombre infini de fois.
Si un état est **récurrent**, tous les états de sa classe sont aussi récurrents. La récurrence est une propriété de la classe.

1.11. Probabilités des états absorbants

Nous avons vu qu'un état absorbant j est un état récurrent pour lequel $P_{jj} = 1$. Il s'agit des états desquels on ne peut pas échapper une fois qu'on y arrive.
Nous posons maintenant la question suivante : quelle est la probabilité a_j que la chaîne de Markov arrive à un état absorbant étant donné qu'elle a commencé à l'état i ?
Etudions la manière de calculer ces probabilités avec un exemple. Considérez la chaîne de Markov suivante :



Les états 4 et 5 sont tous les deux absorbants. Si la chaîne commençait à un de ces états, elle y resterait pour toujours. D'autre part, les états 1, 2 et 3 sont des états transitoires. Si la chaîne devait commencer à l'un de ces états, elle arriverait à un moment donné soit à l'état 4 et elle y resterait, soir à l'état 5 et elle y resterait aussi.

Supposons que l'état initial est 3. Est-ce que la chaîne va finir à l'état 5 ou à l'état 4 ? On ne peut pas être sûr puisque les deux cas son possibles mais on peut intuitivement penser que c'est plus probable que la chaîne finisse à l'état 5 puisqu'il n'y a qu'un pas entre les états 3 et 5.
Nous allons contrôler si notre intuition est correcte une fois que nous aurons calculé les probabilités pour tous les état initiaux possibles.
Commençons par appeler a_i la probabilité que nous finissions à l'état 4 partant de l'état i. Nous pouvons voir tout de suite que si l'état initial est $i = 4$ (on commence à l'état 4), cette probabilité est égale à 1. Ecrivons cette équation :

$$a_4 = 1$$

Si l'état initial est 5, on n'arrivera jamais à l'état 4. On peut donc écrire

$$a_5 = 0$$

Calculons maintenant la probabilité d'arriver à l'état 4 quand on commence à l'état 2.
Il y a deux façons d'arriver à l'état 4. On peut passer directement de 2 à 4 avec probabilité 0.2 ou on peut passer de 2 à 1 avec probabilité 0.8. Quand la chaîne est à l'état 1, la probabilité de finir à l'état 4 est a_1 (voir la définition de a_i pour vous convaincre).
On peut donc écrire

$$a_2 = 0.2a_4 + 0.8a_1$$

Faisons le même raisonnement pour le cas où la chaîne commencerait à l'état 1. La chaîne pourrait passer à l'état 2 avec probabilité 0.6 et ensuite passer à l'état 4 avec probabilité a_2 ou elle pourrait partir dans l'autre directions, vers l'état 3 avec probabilité 0.4 et ensuit passer à l'état 4 avec probabilité a_3. Nous pouvons donc écrire

$$a_1 = 0.6a_2 + 0.4a_3$$

Calculons de la même façon la probabilité à partir de l'état 3. On peut aller vers l'état 2 avec probabilité 0.3 et ensuite vers l'état 4 avec probabilité a_2 ou bien vers l'état 1 avec probabilité 0.5 et ensuite vers l'état 4 avec probabilité a_1. On peut aussi aller vers l'état 5 avec probabilité 0.2 en ensuite vers l'état 4 avec probabilité a_5. L'équation correspondante est donc

$$a_3 = 0.3a_2 + 0.5a_1 + 0.2a_5$$

Nous avons 5 équations avec 5 inconnues (en réalité le deux premières équations nous donnent directement les probabilités a_4 et a_5.

Nous pouvons résoudre ces équations et on obtient :

$$a_1 = 0.6428$$
$$a_2 = 0.7142$$
$$a_3 = 0.5357$$

Maintenant que nous avons calculé ces probabilités, nous pouvons définir b_i comme la probabilité de finir é l'état absorbant 5 si l'on part de l'état i. On peut refaire la même analyse faite pour l'état 4 mais il est plus facile de réaliser que quand on par de l'état i, on va finir tôt o tard soit à l'état absorbant 4, soit à l'état absorbant 5. One peut donc écrire

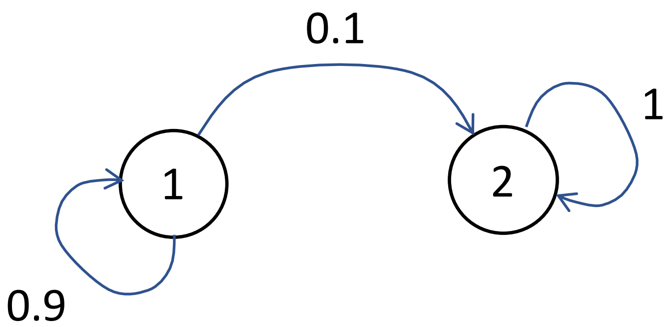
$a_i + b_i = 1$

Cette équation nous permet de trouver les probabilités directement.

1.12. Le nombre de pas pour arriver à un état absorbant

Nous avons vu à la section précédente comment calculer la probabilité d'arriver à un état absorbant dans une chaîne de Markov. Dans cette section, nous allons nous intéresser à une autre variable aléatoire : l'espérance du nombre de pas pour arriver à un état absorbant j depuis un état i.

Appelons μ_i cette valeur espéré du nombre de pas depuis l'état i. Etudions le problème encore une fois par l'intermédiaire d'exemples. Commencer par une chaîne de Markov simple montrée à la figure ci-dessous :



On voit que l'état 1 est absorbant puisqu'une fois dans cette état, la chaîne ne peut plus en sortir. Puisque cette chaîne n'a que deux états, calculons l'espérance du nombre de pas pour arriver à l'état 4 pour le cas où l'état initial est l'état 1 et pour le cas où l'état initial est l'état 2.

Supposons d'abord que la chaîne est à l'état 2. Dans ce cas-ci, il n'y a plus besoin de passer à un autre état et on peut conclure que l'espérance du nombre de pas pour arriver à é'état où la chaîne se trouve déjà est 0. Nous pouvons donc écrire

$\mu_2 = 0$

Voyons maintenant comment calculer l'espérance du nombre de pas jusqu'à l'état absorbant 2 si l'état initial est $X_0 = 1$. Nous allons illustrer deux méthodes.

Method 1

Pour cette première méthode, notez que c'est possible d'arriver à l'état 2 depuis l'état 1 en un pas, en deux pas, en trois pas et ainsi de suite :

$1 \to 2$ (un pas)
$1 \to 1 \to 2$ (deux pas)
$1 \to 1 \to 1 \to 2$ (trois pas)
$\ldots$

Nous pouvons calculer les probabilités de chacun de ces cas :

- la probabilité d'aller directement de $1 \to 2$ en un pas est égale à 0.1 selon le diagramme de transitions de la figure. Nous pouvons illustrer ceci de la manière suivante :

$$1 \overset{0.1}{\to} 2$$

- la probabilité d'arriver en deux pas (le cas $1 \to 1 \to 2$) es la probabilité de passer de l'état 1 à l'état 1 une fois et ensuite de passer de l'état 1 à l'état 2. Les probabilités de chacun des ces pas peuvent être illustrées comme suit :

$$1 \overset{0.9}{\to} 1 \overset{0.1}{\to} 2$$

Cette probabilité est donc le produit 0.9×0.1.

- la probabilité qui correspond à l'arrivé en trois pas ($1 \to 1 \to 1 \to 2$) peut être calculée en regardant l'illustration suivante :

$$1 \overset{0.9}{\to} 1 \overset{0.9}{\to} 1 \overset{0.1}{\to} 2$$

Cette probabilité est donnée par le produit $0.9^2 \times 0.1$

Nous pouvons maintenant généraliser ceci pour calculer la probabilité p_n d'arriver à l'état 2 en n pas :

$$p_n = 0.9^{n-1} \times 0.1$$

Nous pouvons maintenant utiliser la définition de l'espérance d'une variable aléatoire, qui est la somme des produits des valeurs de la variable aléatoire par leur probabilité. La variable aléatoire dans notre cas est le nombre de pas. L'espérance est donc

$$\mu_1 = \sum_{n=1}^{\infty} n p_n = \sum_{n=1}^{\infty} n 0.9^{n-1} \times 0.1$$

Nous pouvons sortir le facteur 0.1

$$\mu_1 = 0.1 \sum_{n=1}^{\infty} n 0.9^{n-1}$$

Le côté droit contient une série infinie de la forme $\sum_{n=1}^{\infty} n x^{n-1}$ dont la somme est connue :

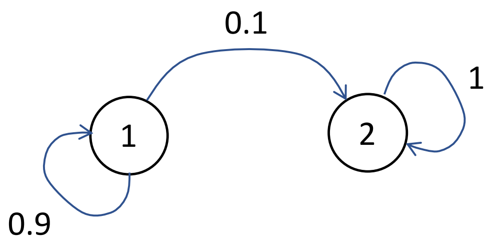
$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}$$

Puisque $x = 0.9$ et que cette somme infinie est multipliée par 0.1, nous pouvons écrire

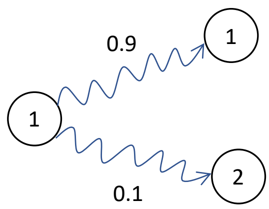
$$\boxed{\mu_1 = 0.1 \times \frac{1}{((1-0.9)^2} = 10}$$

Method 2

Utilisons maintenant une deuxième méthode qui est un peu plus rapide. La chaîne est dessinez encore une fois ci-dessous pour pouvoir y faire référence plus facilement pendant cette explication :



Nous supposons à nouveau que la chaîne commence à l'état 1. Si l'on regarde le diagramme de transitions, on voit que si on attend un coup d'horloge, la chaîne aura soit bouclé sur le même état 1 avec une probabilité 0.9, soit elle est allée à l'état 2 avec probabilité 0.1. Représentons ceci avec une figure.



Que l'on ait fini après ce premier pas à l'état 1 ou à l'état 2, nous aurons déjà fait un pas. Tenez ce fait en tête puisque nous allons l'utiliser un peu plus loin.
Considérons maintenant la situation après ce premier pas. Si la chaîne vient d'atterrir à l'état 1, la probabilité d'arriver depuis là à l'état 2 sera μ_1 (voir la définition de μ_i) et si elle est à atterri à l'état 2, la probabilité d'arriver à l'état 2 sera μ_2. Résumons le raisonnement pour ensuite essayer de le mettre sous la forme d'une équation.
Nous partons de l'état 1 et nous avons seulement deux possibilités : soit nous allons vers l'état 1 avec probabilité 0.9 et après ce premier pas nous allons à l'état 2 avec probabilité μ_1, soit nous allons vers l'état 2 avec probabilité 0.1 et après ce premier pas nous allons à l'état 2 avec probabilité μ_2. Le nombre de pas sera donc soit $(1+\mu_1)$ avec probabilité 0.9, soit $(1+\mu_2)$ avec probabilité 0.1. L'espérance est donc la somme des produits de ces probabilité par le nombre de pas dans chaque cas :
$\mu_1 = 0.9(1+\mu_1) + 0.1(1+\mu_2)$

Nous pouvons maintenant mettre en évidence μ_1 pour obtenir

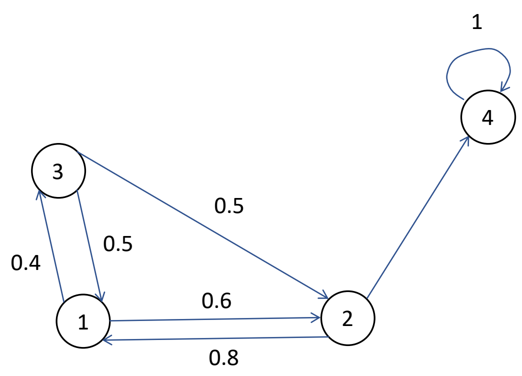
$$\mu_1 = \frac{1 + 0.1\mu_2}{0.1}$$

Puisque l'espérance du nombre de pas pour arriver é l'état 2 depuis l'état 2 est égale à 0 (nous avions mentionné ceci aussi lors du calcul avec la première méthode), nous pouvons donc calculer la valeur de μ_1 :

$$\mu_1 = \frac{1 + 0.1 \times 0}{0.1} = 10$$

qui est la même valeur que nous avions trouvé utilisant la première méthode.

C'est cette deuxième méthode que nous allons apprendre à utiliser de manière systématique pour des cas plus généraux. Nous le ferons aussi en utilisant un exemple.
Considérez la chaîne de Markov dont le diagramme de transitions est ci-dessous. Il y a un état absorbant, l'état 4.



L'espérance du nombre de pas pour arriver à chacun de ces états dépend évidemment de l'état initial. Nous allons calculer l'espérance du nombre de pas pour arriver à l'état 4 pour tous les états initiaux.

<u>Commençons par le cas où la chaîne se trouve à l'état 1.</u>

Depuis l'état 1, la chaîne peut partir
- soit vers l'état 3 avec probabilité 0.4
- soit vers l'état 2 avec probabilité 0.6.

Dans le premier cas, l'espérance du nombre de pas jusqu'à l'état 4 sera alors la somme de 1 (le premier pas jusqu'à l'état 3) et l'espérance μ_3 depuis l'état 3, c'est à dire $(1 + \mu_3)$.

Dans le deuxième cas, l'espérance du nombre de pas jusqu'à l'état 4 est la somme de 1 (le premier pas jusqu'à l'état 2) et l'espérance μ_2 depuis l'état 2, c'est à dire $(1 + \mu_2)$.

Nous pouvons donc écrire
$\mu_1 = 0.4(1 + \mu_3) + 0.6(1 + \mu_2)$

que nous pouvons réécrire ainsi

$$\boxed{\mu_1 - 0.6\mu_2 - 0.4\mu_3 = 1}$$

<u>Regardons maintenant le cas où la chaîne se trouve à l'état 2.</u>

Depuis l'état 2, la chaîne peut partir
- soit vers l'état 1 avec probabilité 0.8
- soit vers l'état 4 avec probabilité 0.2 (cette dernière probabilité n'a pas été écrite explicitement dans le diagramme de transitions mais elle fait que la somme des probabilités sortant de l'état 2 soit 1).

Dans le premier cas, l'espérance du nombre de pas jusqu'à l'état 4 sera alors la somme de 1 (le premier pas jusqu'à l'état 1) et l'espérance μ_1 depuis l'état 1, c'est à dire $(1 + \mu_1)$.

Dans le deuxième cas, l'espérance du nombre de pas jusqu'à l'état 4 est la somme de 1 (le premier pas jusqu'à l'état 4) et l'espérance μ_4 depuis l'état 4, c'est à dire $(1 + \mu_4)$.

Nous pouvons donc écrire
$\mu_2 = 0.8(1 + \mu_1) + 0.2(1 + \mu_4)$
que nous pouvons réécrire ainsi

$$\boxed{\mu_2 - 0.8\mu_1 - 0.2\mu_4 = 1}$$

<u>Si la chaîne se trouve initialement à l'état 3.</u>

Depuis l'état 3, la chaîne peut partir
- soit vers l'état 1 avec probabilité 0.5
- soit vers l'état 2 avec probabilité 0.5.

Dans le premier cas, l'espérance du nombre de pas jusqu'à l'état 4 sera alors la somme de 1 (le premier pas jusqu'à l'état 1) et l'espérance μ_1 depuis l'état 1, c'est à dire $(1 + \mu_1)$.

Dans le deuxième cas, l'espérance du nombre de pas jusqu'à l'état 4 est la somme de 1 (le premier pas jusqu'à l'état 2) et l'espérance μ_2 depuis l'état 2, c'est à dire $(1 + \mu_2)$.

Nous pouvons donc écrire
$\mu_3 = 0.5(1 + \mu_1) + 0.5(1 + \mu_2)$

que nous pouvons réécrire ainsi

$$\boxed{\mu_3 - 0.5\mu_1 - 0.5\mu_2 = 1}$$

Si la chaîne se trouve initialement à l'état 4.

Depuis l'état 4, l'espérance du nombre de pas jusqu'à l'état 4 est égale à 0 puisque la chaîne se trouve déjà à l'état cible. Nous pouvons donc écrire

$$\boxed{\mu_4 = 0}$$

Réécrivons toutes les équations ensemble

$$\mu_1 - 0.6\mu_2 - 0.4\mu_3 = 1$$
$$\mu_2 - 0.8\mu_1 - 0.2\mu_4 = 1$$
$$\mu_3 - 0.5\mu_1 - 0.5\mu_2 = 1$$
$$\mu_4 = 0$$

Remplaçant $\mu_4 = 0$ dans la deuxième équation,

$$\mu_1 - 0.6\mu_2 - 0.4\mu_3 = 1$$
$$\mu_2 - 0.8\mu_1 = 1$$
$$\mu_3 - 0.5\mu_1 - 0.3\mu_2 = 1$$

Ces trois équations représentent un système d'équations avec trois inconnues. IL peut être résolu. Vous pouvez le faire. Vous devriez obtenir les résultats suivants :

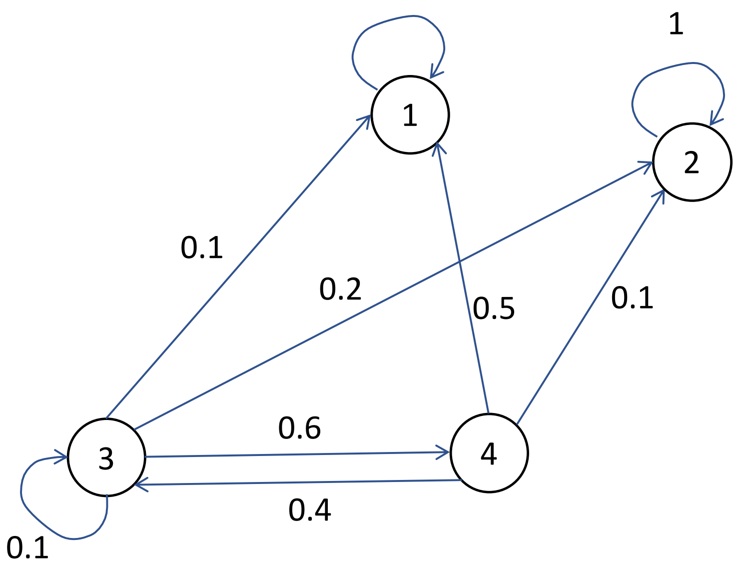
$$\mu_1 = \frac{55}{4}$$

$$\mu_2 = 12$$

$$\mu_3 = \frac{111}{8}$$

Exercices 3

1) Pour la chaîne de Markov suivante :

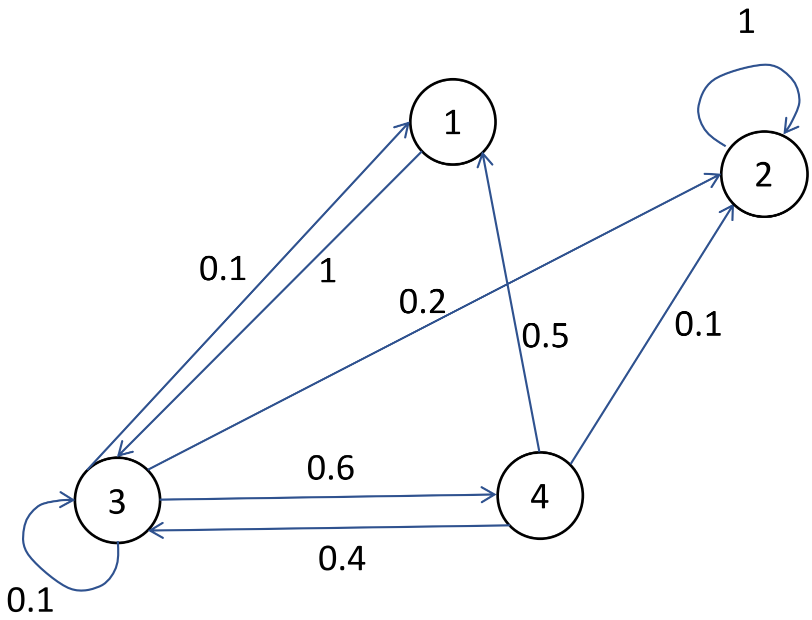


Calculez les probabilités que la chaîne finisse à l'état absorbant 2 pour chacun des cas initiaux suivant : qu'elle commence l'état 3 ou qu'elle commence à l'état 4.

2) Répétez l'exercice précédent calculant cette fois-ci les probabilités que la chaîne finisse à l'état absorbant 1 pour chacun des cas initiaux suivant : qu'elle commence l'état 3 ou qu'elle commence à l'état 4.

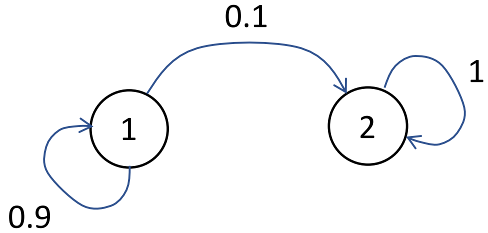
3) Dans les exercices 1) et 2) de cette série, quelle est la probabilité que la chaîne finisse à l'état 1 si elle commence à l'état 2 ?

4) Calculez l’espérance du nombre de pas pour arriver depuis chacun des états de la chaîne suivante jusqu’à l’état absorbant.



5) Sans faire des calculs, pouvez-vous déterminez quelle est l’espérance du nombre de pas pour arriver depuis chacun des états de la chaîne de Markov de l’exercice 1) de cette série jusqu’à l’état absorbant 2 ?

6) Simulez utilisant python la chaine suivante et regardez ce que vous trouvez comme espérance du nombre de pas pour arriver de l’état 1 à l’état 2 (nous avons déjà fait le calcul théorique pour cette chaîne). Obtenez-vous le bon résultat ?



pour cette même chaîne). Obtenez-vous de résultats