Chaînes de Markov

Florence Perronnin

Évaluation de Performances

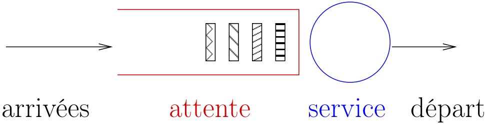
February 27, 2018

Si les gens ne croient pas que les mathématiques sont simples, c'est uniquement parce qu'ils ne réalisent pas à quel point la vie est compliquée.

John von Neumann.

Outline

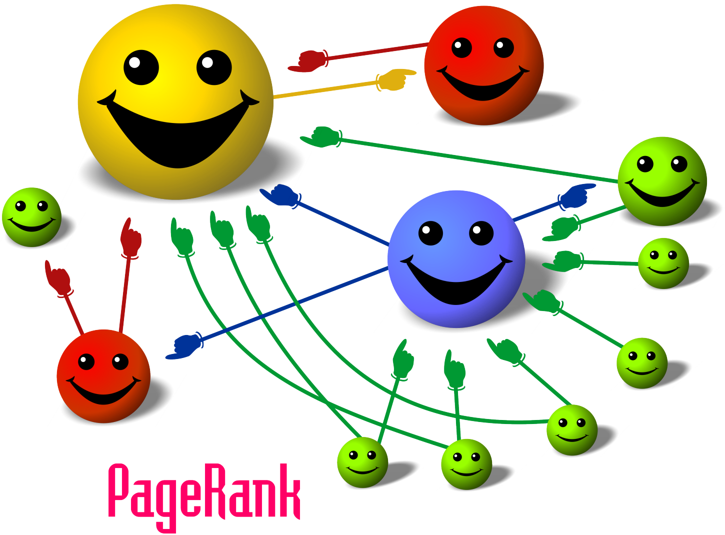
M/M/1



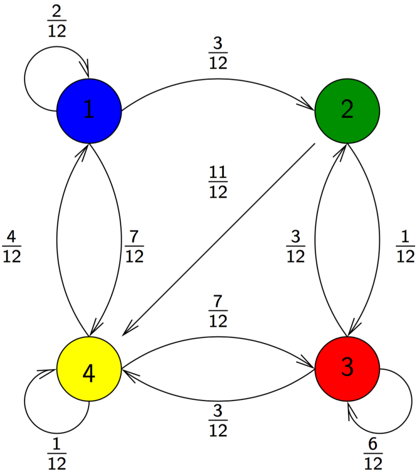
- Si on regardait, pour simplifier, le système une fois par milliseconde? (par exemple)
  - P_a proba d'arrivée à chaque slot
  - P_s proba de fin de service (si au moins un client est présent)
- Et si on regardait le système aux instants de changement d'état ?

On y reviendra plus tard...

PageRank



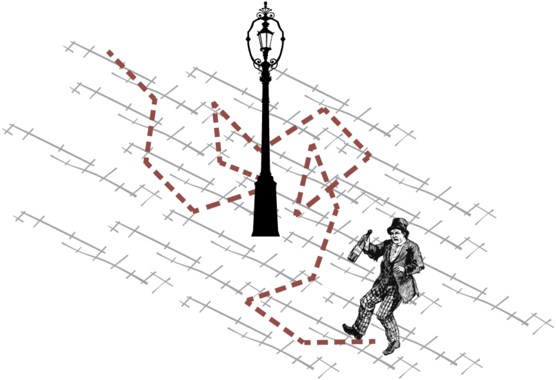
Une version plus petite



Tom et Jerry

Marche aléatoire I

Marche aléatoire II



Autres exemples?

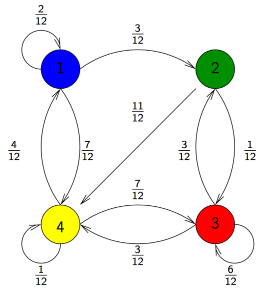
- Shifumi avec une mémoire à 1 coup
- Ou deux...
- M/M/1 en temps discret
- chaîne incluse
- d'autres exemples?

Pour chaque exemple :

- État du système?
- Temps discret?
- Dynamique?
- Graphe de transition?

Outline

(Discrete-time) Markov chains



- État
- Temps discret
- Transitions probabilistes
- État (ou distribution) initiale
- fonction de transition

Transition matrix

$$\frac{1}{12}\begin{pmatrix} 2 & 3 & 0 & 7 \\ 0 & 0 & 1 & 11 \\ 0 & 3 & 6 & 3 \\ 4 & 0 & 7 & 1 \end{pmatrix}$$

Random mapping

- X_n is what we want
- Evolves in state space $\mathcal{S}$
- Trajectory given by $X_0 = x_0$ and

$$X_{n+1} = \Phi(X_n, \xi_n)$$

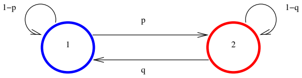
On va chercher à dire des choses sur cet état, qui pourtant change tout le temps...

Outline

Automate Flip-Flop (Bascule)

système ON-OFF

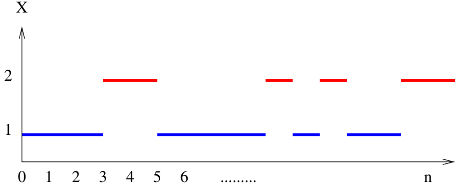
Modèle à deux états :
- ligne de communication
- activité processeur
- ...



Paramètres :
- proportion des transitions : p, q
- Temps de séjour moyen en 1 : $\frac{1}{p}$
- Temps de séjour moyen en 2 : $\frac{1}{q}$

Trajectoire

X_n état de l'automate à l'instant n.



Distribution transitoire

$$\pi_n(1) = \mathbb{P}[X_n = 1] ;$$

$$\pi_n(2) = \mathbb{P}[X_n = 2]$$

Problème

Estimation de π_n : prévision de l'état

Calcul de $\lim_{n \to \infty} \pi_n$: utilisation de ressource

Modèle Mathématique

Probabilités de transition

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

$$\begin{aligned} \mathbb{P}[X_{n+1} = 1 | X_n = 1] &= 1-p \\ \mathbb{P}[X_{n+1} = 2 | X_n = 1] &= p \\ \mathbb{P}[X_{n+1} = 1 | X_n = 2] &= q \\ \mathbb{P}[X_{n+1} = 2 | X_n = 2] &= 1-q \end{aligned}$$

$$\begin{cases} \pi_{n+1}(1) = \pi_n(1)(1-p) + \pi_n(2)q \\ \pi_{n+1}(2) = \pi_n(1)p + \pi_n(2)(1-q) \end{cases}$$

$$\pi_{n+1} = \pi_n P$$

Iterations linéaires

Spectre de P (valeurs propres)

$\mathcal{S}p = \{1, 1-p-q\}$

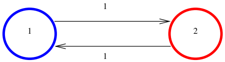
Résolution du système

$|1-p-q| < 1$ **cas non pathologique**

$$\begin{cases} \pi_n(1) = \frac{q}{p+q} + \left(\pi_0(1) - \frac{q}{p+q}\right)(1-p-q)^n; \\ \pi_n(2) = \frac{p}{p+q} + \left(\pi_0(2) - \frac{p}{p+q}\right)(1-p-q)^n; \end{cases}$$

$1-p-q = 1$ $p = q = 0$ **Comportement réductible**

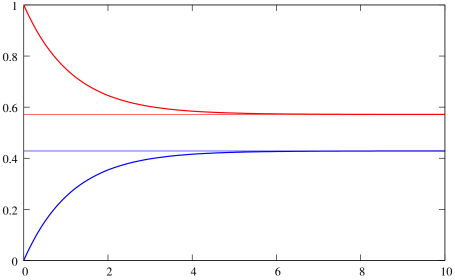
$1-p-q = -1$ $p = q = -1$ **Comportement Périodique**



Comportement récurrent

Exemple numérique

$p = \frac{1}{4}$, $q = \frac{1}{3}$



Convergence rapide
(taux exponentiel)

Comportement stationnaire

$$\begin{cases} \pi_\infty(1) = \frac{q}{p+q}; \\ \pi_\infty(2) = \frac{p}{p+q}. \end{cases}$$

π_∞ unique vecteur probabilité solution

$$\pi_\infty = \pi_\infty P.$$

Si $\pi_0 = \pi_\infty$ alors $\pi_n = \pi_\infty$ pour tout n
Comportement stationnaire

Outline

Chaîne de Markov à temps discret (CMTD)

Propriété de Markov

- Espace d'états discret $\mathcal{S} = \{1, 2, \cdots, K\}$
- Le processus stochastique $\{X_n\}_{n\in\mathbb{N}}$ est une **chaîne de Markov** si

$$\mathbb{P}[X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \cdots, X_0 = i_0] = \mathbb{P}[X_{n+1} = j | X_n = i].$$

pour tout instant n et tout (n+2)-uplet d'états $(i_0, \cdots, i_{n-1}, i, j) \in \mathcal{S}^{n+2}$.
Propriété sans mémoire (conditionnellement à l'état courant)

- La chaîne de Markov est **homogène** si les probabilités de transitions ne dépendent pas du temps :

$$\mathbb{P}[X_{n+1} = j | X_n = i] = \mathbb{P}[X_1 = j | X_0 = i] = p_{i,j}.$$

Interprétation algébrique

La **matrice de transition** $P = ((p_{i,j}))$ est une matrice "stochastique":

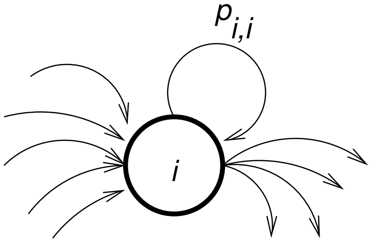
$$p_{i,j} \geq 0;$$

et

$$\sum_j p_{i,j} = 1.$$

Temps de séjour dans un état

Soit T_i temps de séjour dans l'état i :



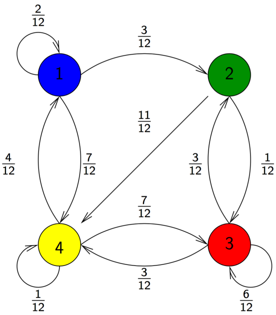
Distribution géométrique $\mathcal{G}(p_{i,i})$:

$$\begin{aligned} \mathbb{P}[T_i = k] &= (1 - p_{i,i})\, p_{i,i}^{k-1} \text{ for } k = 0, 1, \dots \\ \mathbb{E}[T_i] &= \frac{1}{1 - p_{i,i}} \end{aligned}$$

Representation

Transition probabilities

The Markov chain stochastic behaviour can be represented using a transition graph or diagram



Labels on arcs represent transition probabilities

Equivalently, it can be described using the corresponding transition matrix

$$P = \begin{pmatrix} \frac{2}{12} & \frac{3}{12} & 0 & \frac{7}{12} \\ 0 & 0 & \frac{1}{12} & \frac{11}{12} \\ 0 & \frac{3}{12} & \frac{6}{12} & \frac{3}{12} \\ \frac{4}{12} & 0 & \frac{7}{12} & \frac{1}{12} \end{pmatrix}$$

where $p_{i,j} = \mathbb{P}[X_{n+1} = j \mid X_n = i]$

- each line i represents the outgoing transition probabilities from state i
- each column j represents the incoming transition probabilities in state j

Représentations

- État : X_n
- Espace d'états $\mathcal{S}$ (state space)
- Probabilités d'état $\pi_n(i) = \mathbb{P}[X_n = i]$
- $\pi_\mathbf{n}$ sous forme vectorielle : distribution transitoire
- transition en 1 étape (loi des probabilités totales) :

$$\pi_{\mathbf{n+1}} = \pi_\mathbf{n} \mathbf{P}$$

- fonction de transition :

$$X_{n+1} = \Phi(X_n, \xi_n)$$

où

$$\begin{aligned} \Phi : \mathcal{E} \times (0,1) &\to \mathcal{E} \\ X_n &\mapsto \Phi(X_n, \xi_n) \end{aligned}$$

et où ξ_n est l'innovation aléatoire.

Simulation du régime transitoire

1. On part d'un état initial X_0.
   N.B : cet état peut aussi être aléatoire en fixant une distribution initiale.
2. On fixe un critère d'arrêt: par exemple la *durée* maximale T de simulation
3. Pour chaque itération n

   on tire l'innovation aléatoire ξ_n et on applique la fonction de transition Φ.
   (On tire l'événement suivant selon sa probabilité d'occurrence et on applique le changement d'état correspondant.)

Pièges

- Pertinence du critère d'arrêt?
- Choix (et influence) de l'état initial?
- Que peut-on conclure des résultats de cette simulation?
- Combien de trajectoires faut-il simuler?

Chaîne de Markov à temps discret

Equation de Chapman-Kolmogorov

Interprétation trajectorielle

$$\mathbb{P}[X_{m+n} = j|X_0 = i] = \sum_k \mathbb{P}[X_{m+n} = j|X_m = k] \,.\mathbb{P}[X_m = k|X_0 = i]$$

$$= \sum_k \mathbb{P}[X_n = j|X_0 = k] \,.\mathbb{P}[X_m = k|X_0 = i]$$

$$== \sum_k P_{kj}^{(n)} P_{ik}^{(m)}.$$

Probabilité d'une trajectoire $(i_0 \to i_1 \to \ldots \to i_n)$

$$\mathbb{P}[X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_0 = i_0] = \mathbb{P}[X_0 = i_0]\, p_{i_0,i_1}.p_{i_1,i_2} \ldots p_{i_{n-1},i_n}$$

transition en n étapes

$$\mathbb{P}[X_n = j|X_0 = i] = p_{i,j}^{(n)}.$$

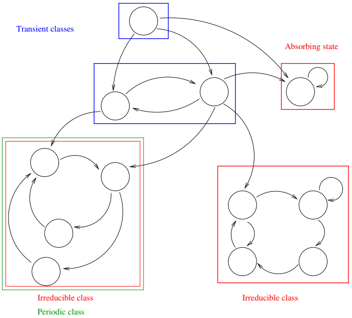
Itération $\Longrightarrow$ produit matriciel

$$((p_{i,j}^{(n)})) = P^n$$

Outline

Classification des états

Analyse de Graphe



Classe irréductible

Composantes connexes
i et j sont dans la même composante (communiquent) s'il existe un chemin de i vers j et un chemin de j vers i avec une probabilité positive.
Les états des classes irréductibles sont appelés **récurrent**

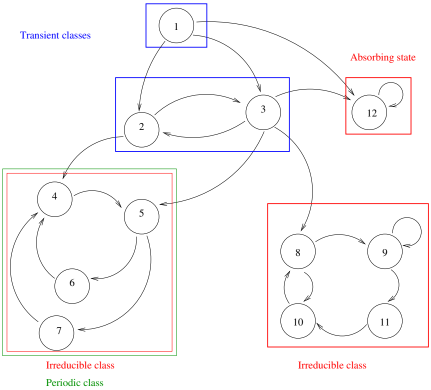
Les autres états sont appelés **transient**

Périodicité

Une classe irréductible est apériodique ssi le PGCD de la longueur de tous ses cycles est 1

Une chaîne de Markov est irréductible s'il existe une unique classe irréductible.
Tout état est donc atteignable depuis tout autre par un *chemin de probabilité non nulle.*

Classification d’états: forme matricielle



	1	2	3	4	5	6	7	8	9	10	11	12
1		*	*									*
2			*									
3		*			*			*				*
4					*							
5						*	*					
6				*								
7				*								
8									*	*		
9									*		*	
10								*				
11										*		
12												1

	1	2	3	4	5	6	7	8	9	10	11	12
1		*	*									*
2			*	*								
3		*			*			*				*
4					1							
5						*	*					

Théorème de Convergence

Soit $\{X_n\}_{n\in\mathbb{N}}$ une chaîne de Markov **homogène irréductible** et **apériodique** de matrice de transition P

1. **Convergence en loi** La distribution transitoire π_n *converge* vers une distribution limite π :

$$\lim_{n\to\infty} \mathbb{P}[X_n = j | X_0 = i] \stackrel{def}{=} \pi(j),$$

2. **Equation d'équilibre** π est l'**unique** vecteur de probabilités solution du système linéaire

$$\pi = \pi P$$

   π est la distribution stationnaire (point fixe)

3. **Vitesse de Convergence**

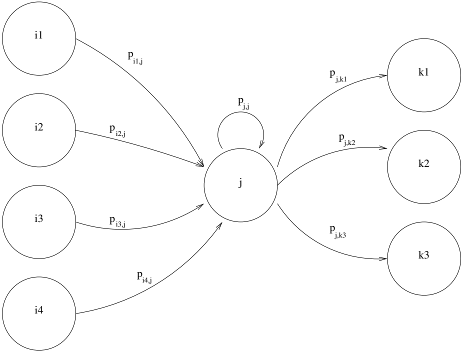
$$||\pi - \pi_n|| \leq C.\alpha_1^n,$$

   α_1 deuxième valeur propre de P et constante C

Interprétation

Équations d'équilibre (*balance equation*) sous forme développée :

$$(\text{taux sortant de } j) \qquad \pi(j) = \sum_i \pi(i) p_{i,j} \qquad (\text{taux entrant dans } j)$$



Si $\pi_0 = \pi$ le processus est stationnaire ($\pi_n = \pi$)

Normalisation

$$\sum_{i \in \mathcal{S}} \pi(i) = 1$$

Ergodicité

Soit $\{X_n\}_{n\in\mathbb{N}}$ une chaîne de Markov **homogène irréductible** et **apériodique** de matrice de transition P

$T_n(i)$: temps passé dans l'état i pour la trajectoire $X_0, X_1, \cdots, X_{n-1}$

$$T_n(i) = \sum_{k=0}^{n-1} \mathbb{1}_{[X_k=i]}$$

Théorème Ergodique

$$\text{Time average} \qquad \lim_{n\to\infty} \frac{1}{n} T_n(i) = \pi(i). \qquad \text{Ensemble average}$$

Pour toute fonction de gain f

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \mathbb{E}_\pi f(X).$$

Gain moyen par unité de temps

Outline

Repair shop revisited

Change the ON/OFF model: assuming after 3 days of unavailability, the resource is replaced by a new one.
Question:

1. what is the new transition diagram?

Umbrellas

- 2 umbrellas for commuting (home $\leftrightarrow$ work).
- When it rains: if an umbrella is available, take it.
- When it doesn't rain: don't take an umbrella.
- $\mathbb{P}[\text{rain}] = p$.

Questions:

1. State? State space?
2. Transition diagram / matrix?
3. What is the limiting probability of getting soaked?

Modeling the umbrella problem # 1

We can model the problem by considering the position of the umbrellas:

$$X_n = (u_n, v_n)$$

where u_n (resp. v_n) $\in \{H, W\}$ (for "home"and "work") denotes the position of the first (resp. second) umbrella.
However...
X_n is not a Markov chain, because we still need to know where the person is.

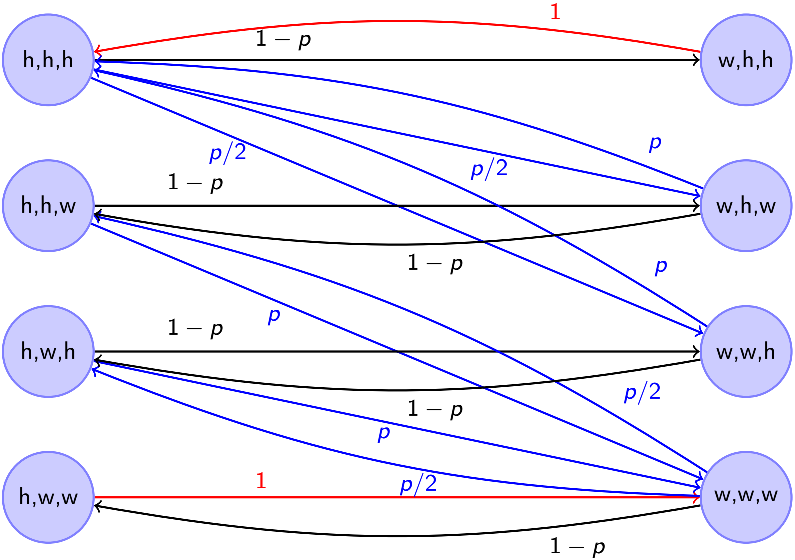
Solution # 1

We can enrich the system state with the position of the person:

$$X_n = (x_n, u_n, v_n)$$

The system is now a Markov chain, with state space: $\{H, W\} \times \{H, W\} \times \{H, W\}$ (8 possible states).

Modeling the umbrella problem # 1



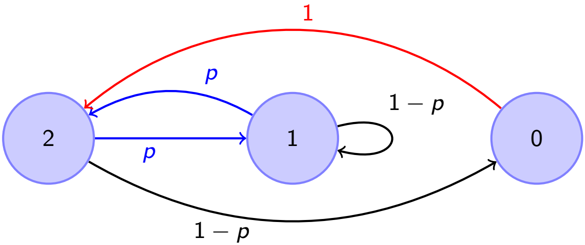
Modeling the umbrella problem # 2

The above model is completely symmetric and very redundant. Actually, we do not need to know the exact place of the person if we know how many umbrellas are available with her.
Let us define a new model:

$$X_n = (\text{number of available umbrellas at the place where the owner is}) \in \{0, 1, 2\}$$

Actually X_n is a Markov chain because the number of umbrellas at the other place directly depends on the number of umbrellas at the previous place... and on the chance of rain, which is independent of the umbrella location.
The new state space is considerably smaller (3 states).
In addition, we want to know the chances of getting soaked, i.e. the probability of having no umbrella (state 0) when it rains.

Modeling the umbrella problem # 2



The DTMC is clearly irreducible (all states communicate), aperiodic (loop on state 1) and finite-state. So there is a limiting distribution π which satisfies the following equations:

$$
\begin{array}{lcl}
\pi(0) & = & (1-p)\,\pi(2) \\
\pi(1) & = & (1-p)\,\pi(1) + p\,\pi(2) \\
\pi(2) & = & \pi(0) + p\,\pi(1)
\end{array}
$$

Solving the umbrella problem

The above equations imply that :

$$\begin{aligned} \pi(2) &= \frac{1}{1-p}\pi(0) \\ \pi(1) &= \pi(2) \end{aligned}$$

We can now use the fact that:

$$\pi(0) + \pi(1) + \pi(2) = 1$$

to get :

$$\pi(0) = \frac{1-p}{3-p}, \qquad \pi(1) = \pi(2) = \frac{1}{3-p}$$

So the limiting probability of having no umbrella is $\frac{1-p}{3-p}$.
The probability of getting soaked is therefore:

$$\begin{aligned} \mathbb{P}[\text{rain and state } 0] &= \mathbb{P}[\text{rain}]\,\mathbb{P}[\text{state } 0] \quad \text{by independence} \\ &= p\pi(0) \\ &= \frac{p(1-p)}{3-p} \end{aligned}$$

Transmission avec une probabilité d'erreur

Un routeur reçoit des paquets de taille identique arrivant dans les intervalles de temps disjoints. Hypothèses:

- maximum 1 paquet par unité de temps $]t(n), t(n+1)[$
- arrivées **i.i.d**
- buffer infini
- erreur de transmission avec proba $(1-p)$ (occurrences **i.i.d**), indépendamment des arrivées
- durée de transmission: 1 unité de temps
- début à l'instant t(n) si au moins 1 paquet en attente

Modélisation

- $A(n) \in \{0,1\}$ nombre d'arrivées dans $]t(n), t(n+1)[$: $\sim \mathcal{B}(a)$
- $D(n)$ nombre de transmissions réussies dans $]t(n), t(n+1)[$: $D(n) \sim \mathcal{B}(p)$
- $X(n)$ nombre de paquets dans le routeur

équations d'évolution:

$$X(n+1) = \begin{cases} A(n) & \text{si } X(n) = 0 \\ X(n) + A(n) - D(n) & \text{si } X(n) > 0 \end{cases}$$

Résolution

1. Preuve que $X(n)$ est une CMTD
2. Irréductibilité
3. Apériodicité
4. Espace d'état infini : il faut déterminer si $\pi P = \pi$ admet une unique solution telle que $\sum_{i=0}^{\infty} \pi(i) = 1$
5. Taux d'utilisation? $1 - \pi(0)$
6. Débit? $p(1 - \pi(0))$

Caractérisation de la CMTD

$$P = \begin{pmatrix} 1-a & a & 0 & \dots \\ (1-a)p & ap+(1-a)(1-p) & a(1-p) & 0... \\ 0 & (1-a)p & ap+(1-a)(1-p) & a(1-p)... \\ 0 & \ddots & \ddots & \ddots \end{pmatrix}$$

Irréductibilité

On veut prouver que tous les états communiquent. Formellement:
$\forall (i,j), \exists (n,k) \in \mathbb{N}^2 \,/\, \mathbb{P}[X_{n+k} = j | X_n = i] > 0$
Il suffit donc d'exhiber, pour tout i et j, au moins 1 chemin de probabilité non nulle de i vers j.

- Cas $i < j$: prenons $k = j - i$: il suffit d'avoir k arrivées consécutives

$$\mathbb{P}[X_{n+k} = j | X_n = i] = [a(1-p)]^k > 0$$

- Cas $i > j$: prenons $k = i - j$: il suffit d'avoir k départs consécutifs

$$\mathbb{P}[X_{n+k} = j | X_n = i] = [p(1-a)]^k > 0$$

- Cas $i = j > 0$ Prenons $k = 1$,

$$\mathbb{P}[X_{n+1} = i | X_n = i] = ap + (1-a)(1-p) > 0$$

- Cas $i = j = 0$ Prenons $k = 1$,

$$\mathbb{P}[X_{n+1} = 0 | X_n = 0] = 1 - a > 0$$

Résolution des équations d'équilibre $\pi P = \pi$

$$\begin{aligned}
\pi(0) &= (1-a)\pi(0) + (1-a)p\pi(1) && (1)\\
\pi(1) &= a\pi(0) + [ap + (1-a)(1-p)]\pi(1) + (1-a)p\pi(2) && (2)\\
\vdots &= \vdots \\
\pi(i) &= a(1-p)\pi(i-1) + [ap + (1-a)(1-p)]\pi(i) + (1-a)p\pi(i+1) && (3)
\end{aligned}$$

Les deux premières équations sont un peu différentes de toutes les autres. On exprime tout en fonction de $\pi(0)$ (que l'on déterminera par l'équation de normalisation):

$$\begin{aligned}
\pi(1) &= \frac{a}{p(1-a)}\pi(0) \qquad \text{d'après (1)} && (4)\\
\pi(2) &= \frac{a(1-p)}{[p(1-a)]^2}\pi(0) \qquad \text{d'après (4) et (2)} && (5)
\end{aligned}$$

De même on peut montrer par récurrence en utilisant (3) que :

$$\pi(i) = \frac{1}{1-p}\left(\frac{a(1-p)}{p(1-a)}\right)^i \pi(0) \quad (i \geq 2)$$

Résolution avec l'équation de normalisation $\sum_{i=0}^{\infty} \pi(i) = 1$

On utilise le fait que π est un vecteur de probabilités pour terminer la résolution.

Stabilité

$\pi(i)$ est une suite géométrique, de raison $\frac{a(1-p)}{p(1-a)}$. La série converge si et seulement si la raison est < 1, *i.e.* ssi $a < p$ (ce qui paraît naturel !).

Lorsque la condition de stabilité est remplie, le système d'équations d'équilibre admet donc une unique solution π qui soit un vecteur de probabilités. Cette solution s'obtient par:

$$
\begin{aligned}
\sum_{i=0}^{\infty} \pi(i) &= 1 && (6) \\
&= \pi(0) + \sum_{i=1}^{\infty} \frac{1}{1-p} \left(\frac{a(1-p)}{p(1-a)} \right)^i \pi(0) && (7) \\
&= \pi(0) \left(1 + \frac{1}{1-p} \left(\frac{1}{1 - \frac{a(1-p)}{p(1-a)}} - 1 \right) \right) && (8)
\end{aligned}
$$

Transmission avec erreurs revisitée

1. M/M/1 observée à intervalles fixes :
   - Que devient le diagramme de transition si l'on autorise les arrivées multiples? (loi de Poisson par exemple, de paramètre λ)
   - Et les départs multiples?
2. M/M/1 observée aux instants de saut (chaîne "incluse") :
   - Que devient le diagramme de transition ?
   - Calculer les probabilités de transition en fonction de a et de p.
3. Serveur supplémentaire en cas de saturation:
   - si le nb de client dans la file dépasse un seuil T, on allume un second serveur (de même vitesse), partageant la même file.
   - En-dessous de ce seuil, le second serveur est éteint.
   - Que devient le diagramme de transition?
4. Hysteresis:
   - On n'éteint le second serveur que si la file redescend sous un autre seuil $B < T$. Que devient le processus $X(t)$?
   - Corriger le modèle et donner son diagramme de transition.