TD 4 : Schémas de Matthes et chaînes de Markov finies

lionel.rieg@ens-lyon.fr

1 Schéma de Matthes

Exercice 1 (Commutation de paquets)

Un réseau à commutation de paquets (comme internet) transporte des paquets contenant de l'information. Il est formé de canaux de communication et de routeurs chargés d'aiguiller les paquets sur le bon canal. À chaque fin de transmission d'un paquet, ou bien le routeur qui l'a reçu le fait disparaître dans un réseau local (le paquet est arrivé à destination), ou bien il tente de le transmettre sur un canal. Le débit de transmission nécessaire à un instant à travers un canal donné peut potentiellement excéder sa capacité. Les routeurs sont donc équipés de mémoires pour mettre en attente les paquets excédentaires.

On va supposer pour simplifier que les routeurs ont une mémoire illimitée et qu'ils choisissent instantanément le canal vers lequel rediriger un paquet tout juste arrivé, si bien que les ressources critiques sont les canaux, qu'on va supposer de capacité égale à un. Chaque routeur est relié à un réseau local qui émet et reçoit des paquets. Dans un routeur, l'apparition d'un nouveau paquet qui sera aiguillé sur le canal c est gouvernée par un processus de Poisson de paramètre λ_c. Le routage est représenté par une matrice de transition qui donne la probabilité p_{ij} qu'un paquet sortant du canal i entre dans le canal j. Les paquets sont transmis de manière séquentielle et la durée de transmission d'un paquet suit une loi exponentielle de paramètre μ_c (qui dépend de la fiabilité du canal c).

1. Préciser le devenir d'un paquet (avec les probabilités) qui arrive sur un routeur.
2. Donner le schéma de Matthes correspondant à cette situation. Écrire la procédure de simulation de ce schéma en pseudo-code.

2 Chaînes de Markov

Les chaînes de Markov sont définies par la propriété éponyme que voici :

$$\forall n, \forall x_1, \ldots, x_n,\ \mathbb{P}(X_n = x_n \,|\, X_0 = x_0,\, X_1 = x_1, \ldots\, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = x_n \,|\, X_{n-1} = x_{n-1})$$

Lorsque cette probabilité conditionnelle ne dépend pas de n, la chaîne est dite *homogène*. Cela correspond à son indépendance vis-à-vis du temps. On peut alors représenter les transitions de la chaîne par une multiplication matricielle.

Exercice 2

On lance un dé à six faces. Est-ce que les variables aléatoires suivantes sont des chaînes de Markov ? Si oui, préciser leurs graphes de transition.

1. La plus grande valeur observée jusqu'au n^{e} lancer.
2. Le nombre de six obtenus jusqu'au n^{e} lancer.
3. Le nombre de lancers depuis le dernier six.
4. Le nombre de lancers jusqu'au prochain six.

Peut-on dégager un schéma général ?

Exercice 3

On considère un système composé de n stations qui émettent des messages. Le temps est discret et un entier $W > 0$ est fixé.

Si à un instant une station i à un message à émettre, elle tire indépendamment un entier aléatoire W_i dans $\{0, 1, \ldots, W-1\}$. A chaque nouvel instant, elle décrémente de 1 son entier W_i. A l'instant où W_i vaut 0, elle émet son message.

Si une seule station émet à un instant donné, son message est bien transmis. Si elle a un autre message à émettre, elle tire à l'instant suivant un nouvel entier aléatoire et recommence la procédure.

Si deux stations ou plus émettent en même temps, les messages se brouillent, on parle de *collision*. L'émission de tous ces messages échoue. Chaque station concernée conserve son message et tire à l'instant suivant et indépendamment un nouvel entier aléatoire. La procédure continue ainsi de suite.

1. Que se passe-t-il dans le système si on choisit $W = 1$?
2. On suppose que à l'instant $t = 0$, chaque station a un nombre fixé, fini ou infini, de messages qu'elle doit transmettre les uns après les autres. Montrer que l'on peut décrire la dynamique du système sous forme d'une chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ à valeurs dans un ensemble d'états bien choisi. Illustrer en dessinant le graphe de transition dans le cas $n = 2$, $W = 2$ et 1 message à transmettre par station, ainsi que dans le cas $n = 2$, $W = 3$ et une infinité de messages par station.

Exercice 4

Donner les matrices de transition et les graphes associés pour les problèmes suivants et calculer, si elles existent, les distributions limites en utilisant des relations de récurrences.

1. Doudou, le hamster paresseux, ne connaît que trois endroits dans sa cage : les copeaux où il dort, la mangeoire où il mange et la roue où il fait de l'exercice. Ses journées sont assez semblables les unes aux autres, et son activité se représente aisément par une chaîne de Markov. Toutes les minutes, il peut soit changer d'activité, soit continuer celle qu'il était en train de faire. L'appellation processus sans mémoire n'est pas du tout exagérée pour parler de Doudou.
   – Quand il dort, il a 9 chances sur 10 de ne pas se réveiller la minute suivante.
   – Quand il se réveille, il y a 1 chance sur 2 qu'il aille manger et 1 chance sur 2 qu'il parte faire de l'exercice.
   – Le repas ne dure qu'une minute, après il fait autre chose.
   – Après avoir mangé, il y a 3 chances sur 10 qu'il parte courir dans sa roue, mais surtout 7 chances sur 10 qu'il retourne dormir.
   – Courir est fatigant ; il y a 8 chances sur 10 qu'il retourne dormir au bout d'une minute. Sinon il continue en oubliant qu'il est déjà un peu fatigué.
2. Ruine du joueur : un joueur commence avec un pactole de s €. Il lance une pièce : si elle tombe sur pile il gagne 1 €, sinon il perd 1 €. Il joue jusqu'à atteindre une fortune de M € ou être ruiné.
3. Marche aléatoire isotrope sur $\mathbb{Z}$
4. Diffusion de Ehrenfest : N particules sont enfermées dans deux boîtes reliées par un tube. À chaque instant (discret), une particule est choisie de façon uniforme et passe d'une boîte à l'autre. On s'intéresse au nombre de particules dans chaque boîte.

Exercice 5 (Analyse d'une multinationale)

Une multinationale doit traiter un dossier très sensible. Il est certain que lorsqu'il sera découvert, le cadre qui en est actuellement en charge se fasse remercier pour incompétence, ceci afin de préserver l'image de l'entreprise. Chaque personne cherche donc à se débarrasser du dossier, ce qu'elle peut faire de deux façons :

– ou bien le donner à un autre cadre de son service choisi uniformément (toutes les personnes d'un même service se connaissent),
– ou bien le faire sous-traiter par un autre service dont elle est responsable (le dossier est alors donné à un cadre de ce service, choisi uniformément).

On suppose que la propension de chaque cadre à choisir (s'il en a la possibilité !) l'une de ces deux options est constante. On s'intéresse à la probabilité se faire remercier, c'est-à-dire la probabilité d'avoir le dossier sensible. Comme sa découverte par la presse va prendre du temps, on ne considère que le comportement asymptotique des mouvements du dossier. Voici la liste des cadres de la multinationale :

personne	A	B	C	D	E	F	G	H	I	J	K	L
membre des services	1	1	1	2	2	2,3	3	4	4	5	5	5
responsable des services	1	2 3	3	4			5					
propension à sous-traiter	0,4	0,5 0,5	0,7	0,4	0	0	0,9	0	0	0	0	0

1. Donner le graphe de transition des mouvements du dossier.
2. Y a-t-il des personnes à l'abri ? Si oui, pourquoi ? Comment les caractériser ?
3. Peut-on simplifier la chaîne de Markov ? Si oui, pourquoi ?
4. Donner une présentation matricielle du problème et y répondre (on admet la convergence de la chaîne de Markov).
5. On suppose que chaque service possède une secrétaire. Lors d'un transfert de dossier, le cadre qui gère le dossier choisit à quel service le renvoyer (le sien ou l'un des ceux dont il est responsable) et c'est la secrétaire du service concerné qui fait l'attribution (uniforme) au sein du service. Comment est modifiée la chaîne de Markov ? Quelle nouvelle propriété a-t-elle ? Peut-on adapter l'analyse précédente ?

Solutions

▸ Exercice 1

1. On note i le canal par lequel le paquet arrive. Il peut soit être redirigé vers un autre canal j avec probabilité p_{ij} ou être envoyé dans le réseau local du routeur avec probabilité $1 - \sum_{j \neq i} p_{ij}$.
2. L'ensemble $\mathcal{E}$ des états est le nombre de paquets en attente ou en cours de transmission dans chaque canal. On ignore donc l'existence des routeurs. C'est donc un vecteur ι de $\mathbb{N}^N$ avec N le nombre de canaux. L'ensemble $\mathcal{S}$ des sources est divisées en sources « positives » $\alpha_1, \ldots, \alpha_N$ qui font apparaître (depuis un réseau local) des paquets dans le réseau et sources négatives $\beta_1, \ldots, \beta_N$ qui font disparaître du buffer les paquets qui ont traversés le canal p (elles représentent les transitions). L'ensemble $A(\iota)$ des sources actives est $\{\alpha_1, \ldots, \alpha_N\} \cup \left\{\beta_p \;\middle|\; \iota_p > 0\right\}$: des paquets apparaissent en permanence et seul les canaux non vide effectuent des transferts. Les vitesses sont toutes égales à 1 par hypothèse. les termes non nuls des distributions de transition sont

$$\begin{aligned} p(\alpha_p, \iota, \iota + e_p) &= 1 \\ p(\beta_p, \iota + e_p, \iota + e_q) &= p_{pq} \\ p(\beta_p, \iota + e_p, \iota) &= 1 - \sum_{q=1}^{N} p_{pq} \end{aligned}$$

Les fonctions de répartition qui définissent les distributions des sources sont :

$$\begin{aligned} F_{\alpha_p}(x) &= 1 - \mathrm{e}^{-\lambda_p x} \\ F_{\beta_p}(x) &= 1 - \mathrm{e}^{-\mu_p x} \end{aligned}$$

▸ Exercice 2

Le schéma général est donné par le théorème suivant :

> Si $(Z_n)_{n\in\mathbb{N}}$ est une suite de v. a. i. i. d. à valeur dans F, si $f : E \times F \to E$ est une fonction et si X_0 est une v. a. indépendante des Z_n, alors la suite de v. a. $X_{n+1} = f(X_n, Z_n)$ définit une chaîne de Markov homogène.

Ici, les Z_n sont les lancers successifs du dé, X_0 vaut 0 (sauf dans le dernier cas) et f varie en fonction de ce qu'on compte. Ainsi, les trois premiers exemples proposés sont des chaînes de Markov. Quant au dernier, il ne faut pas se laisser impressionner par la dépendance vers le futur : le temps d'attente ne dépend pas de historique de ces temps. On peut même en détailler la loi : si X_n est non nul, $X_{n+1} = X_n - 1$ et si $X_n = 0$, le temps d'attente suit une loi géométrique de paramètre $\frac{1}{6}$.

Cependant, il est possible de fabriquer à partir de suites i. i. d. des suites qui ne soient pas des chaînes de Markov. Par exemple, Si $(Z_n)_{n\in\mathbb{N}}$ est une suite de v. a. i. i. d. telles que $\mathbb{P}(Z_n = 1) = \mathbb{P}(Z_n = -1) = \frac{1}{2}$, on définit X_n par :

$$X_{2n+1} = Z_n \qquad\qquad X_{2n} = Z_{n-1} Z_n$$

On a alors $\mathbb{P}(X_{2n+1} = 1 \,|\, X_{2n} = -1\,) = \frac{1}{2}$ mais $\mathbb{P}(X_{2n+1} = 1 \,|\, X_{2n} = -1 \wedge X_{2n-1} = 1\,) = 0$.

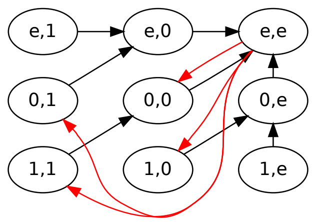
▸ Exercice 3

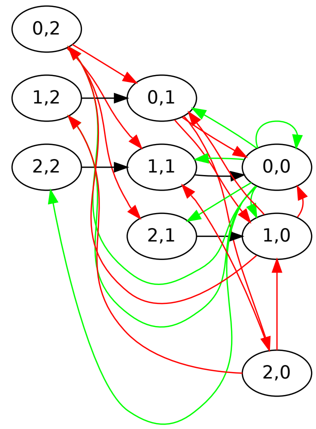
1. Lorsque $W = 1$, en cas de collision, les stations ré-émettent à l'instant d'après (W_i est toujours égal à 0) donc sont à nouveau en collision. Ainsi, la moindre collision bloque tout le système.
2. L'ensemble générique d'états à considérer est : $((\mathbb{N} \cup \{+\infty\} \times W)^S$ où $\mathbb{N} \cup \{+\infty\}$ représente le nombre de messages à transmettre et S est le nombre de stations. Lorsque le nombre de message à transmettre est nul, la station est au repos et il n'est plus nécessaire de conserver des informations pour W_i. En tenant compte de cette observation, l'espace des états devient $(\{\varnothing\} \cup ((\mathbb{N}^* \cup \{+\infty\}) \times W))^S$.

   Pour $n = 2$, $W = 2$ et 1 message à transmettre, il n'est pas nécessaire de stocker le nombre de messages à envoyer donc les états intéressants sont pour chaque station : e qui indique le repos, $W_i = 0$ avec un message à envoyer et $W_i = 1$ avec un message à envoyer. Cela fait donc neuf états au total.

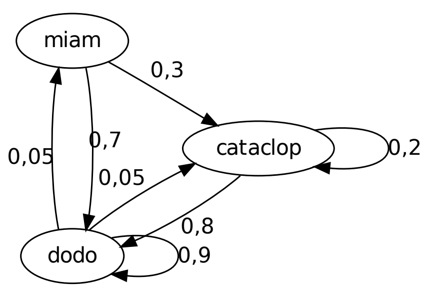
Pour $n = 2$, $W = 3$ et une infinité de messages à envoyer, il n'est pas besoin de stocker le nombres de messages à envoyer et il n'y a pas d'état de repos. Ainsi la seule grandeur intéressante est W_i. On a donc à nouveau 9 états possibles.

Dans les deux figures ci-après, les flèches noires correspondent aux transitions de probabilité 1 alors que les rouges sont uniformes en fonction de leur nombre (quatre ou neuf).



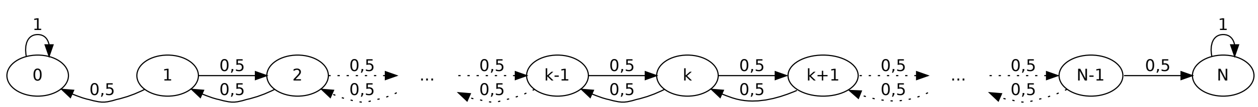


▸ **Exercice 4** La méthode générale pour calculer une distribution limite consiste à résoudre le système suivant : $\pi = P\pi$ où P est la matrice de transition de la chaîne de Markov, ce qui correspond exactement à un système de relations de récurrence. On rappelle que cette dernière n'existe que si la chaîne est homogène. On peut réécrire cette équation sous la forme $\pi(P - I) = 0$.

1. 

$$\begin{pmatrix} 0,9 & 0,7 & 0,8 \\ 0,05 & 0 & 0 \\ 0,05 & 0,3 & 0,2 \end{pmatrix}$$

La distribution stationnaire est $\frac{1}{181}(160, 8, 13)$, qui vaut approximativement $(0,884,\ 0,044,\ 0,72)$. Doudou passe plus de 88% de son temps à dormir ! Il n'y a ici qu'une unique distribution stationnaire.



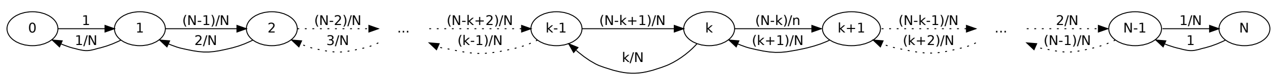
2.
$$\begin{pmatrix} 1 & 0,5 & & & & \\ & 0 & \ddots & & & \\ & 0,5 & \ddots & \ddots & & \\ & & \ddots & \ddots & 0,5 & \\ & & & \ddots & 0 & \\ & & & & 0,5 & 1 \end{pmatrix}$$

Il y a deux distributions stationnaires : la ruine ou la fortune (et toute combinaison convexe des deux). Cela peut se voir facilement en remarquant que ces deux états sont absorbants donc s'il y a une probabilité non nulle d'y arriver, ils absorbent les probabilités des états qui y mènent. Comme la chaîne (privée de ces extrémités) est irréductible, aucun autre état ne peut avoir de probabilité non nulle. Il y a dans ce cas une infinité de distributions stationnaires.

3. C'est la limite du cas précédent lorsque les bornes tendent vers l'infini. On ne peut plus alors parler de matrice (elle serait infinie) mais d'opérateur de transition.

   Par symétrie, tous les états doivent avoir même probabilité. Et de fait, $c(\dots,1,\dots)$ pour $c \in \mathbb{R}^+$ est une mesure stationnaire mais n'est pas une distribution car la condition de somme égale à un est irréalisable. Il n'y a pour cet exemple aucune distribution stationnaire.

4.



Il est légitime d'un point de vue physique de chercher une solution qui soit réversible temporellement. On utilise donc la condition de *réversibilité à l'équilibre* et on résout $\pi_i p_{ij} = \pi_j p_{ji}$. Cette équation n'a d'intérêt que pour $j \in \{i-1, i+1\}$ et par symétrie on peut se limiter à un seul de ces cas. On cherche donc à résoudre $\pi_i \frac{m-i}{m} = \pi_{i+1}\frac{i+1}{m}$. La solution est de la forme $\binom{m}{i}$ à un facteur de normalisation près. En effet :

$$\frac{m-i}{m}\pi_i = \frac{m-i}{m}\binom{m}{i} = \frac{m-i}{m}\frac{m!}{(m-i)!\,i!} = \frac{(m-1)!}{(m-i-1)!\,i!} = \frac{i+1}{m}\frac{m!}{(m-i-1)!\,(i+1)!} = \frac{i+1}{m}\binom{m}{i+1} = \frac{i+1}{m}\pi_{i+1}$$

On peut vérifier (mais c'est inutile car l'équation de réversibilité à l'équilibre entraîne celle d'invariance par la matrice de transition) que c'est bien une mesure stationnaire de la chaîne de Markov :

$$\begin{aligned}\sum_k \pi_k p_{ki} &= \pi_{i-1}\frac{m-i+1}{m} + \pi_{i+1}\frac{i+1}{m} = \binom{m}{i-1}\frac{m-i+1}{m} + \binom{m}{i+1}\frac{i+1}{m} = \frac{(m-1)!}{(m-i)!\,(i-1)!} + \frac{(m-1)!}{(m-i-1)!\,i!}\\ &= (m-1)!\frac{i+(m-i)}{(m-i)!\,i!} = \frac{m!}{(m-i)!\,i!} = \binom{m}{i} = \pi_i\end{aligned}$$

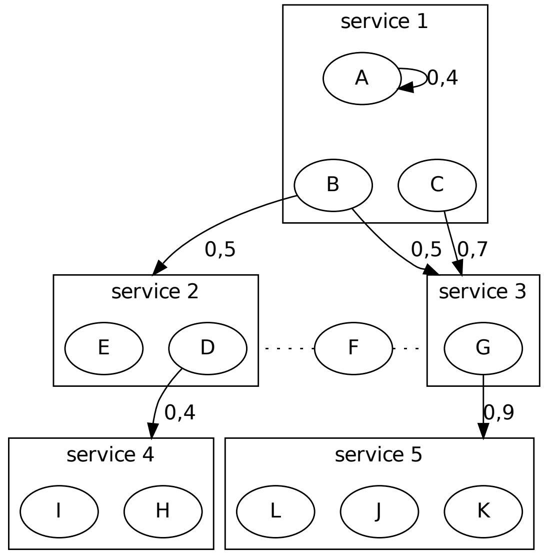
Le coefficient de normalisation est

$$\left(\sum_i \pi_i\right)^{-1} = \left(\sum_i \binom{m}{i}\right)^{-1} = (2^m)^{-1} = 2^{-m}$$

Enfin, on reconnaît que cette distribution est une loi binomiale de paramètres N et $\frac{1}{2}$, *i.e.* celle qu'on aurait obtenu en chaisissant aléatoirement avec une pièce équilibrée dans quelle urne mettre chaque particule.

▸ Exercice 5

1. Malgré le dessin, le noeud F fait partie des services 2 et 3.



2. Les personnes à l'abri sont celles pour lesquelles la probabilité d'une infinité de retour est égale à 0 (*i.e.* la probabilité d'un retour n'est pas égale à 1). Cela correspond aux services 1, 2, et 3. Ces états sont appelés *transcients* par opposition aux états *récurrents*. Lorsque tous les états sont récurrents, la chaînes est dite *irréductible*.
3. L'absence de retour possible dans le graphe est lié à la décomposition en composantes fortement connexes (appelées *classes de communication* dans la théorie des chaîne de Markov). Les personnes sur la sellette sont précisément celles qui appartiennent à une composante fortement connexe finale. On peut donc commencer une étude à haut niveau sur le graphe des composantes fortement connexes (dans lequel on contracte les noeuds d'une même classe de communication) pour déterminer quelle est la probabilité d'arriver dans les différentes classes finales. On fera ensuite une étude plus fine au sein de chaque classe.
4. La distribution limite π_∞ (qu'on suppose exister) est stable par multiplication avec la matrice de transition M car $\pi_\infty = M^\infty \pi_0 = MM^\infty \pi_0 = M\pi_\infty$. On recherche donc un vecteur propre de M *i.e.* un vecteur du noyau de $M - I$.
5. On remarque les secrétaires ont le dossier en main une étape sur deux vu que chaque mouvement passe par elles. Ainsi, on ne peut avoir convergence vers une mesure stationnaire. De plus, tout chemin de retour a une longueur paire. On dit que la chaîne est *périodique* de période 2. Afin de se ramener à une chaîne qui possède une distribution stationnaire, on étudie les chaînes D_{2n} et D_{2n+1}. On retrouve alors des chaînes pour lesquelles le pgcd des longueurs de chemins de retour est 1, qu'on appelle chaîne *apériodique*. En revanche, comme on a modifié la chaîne, on a peut-être perdu la propriété d'irréductibilité pour la matrice M^2, ce qui nous contraint à reprendre la même analyse. Ultimement, on obtient une chaîne irréductible et apériodique pour laquelle un théorème nous donne l'existence d'une mesure stationnaire.