Examen "Chaînes de Markov"

Mercredi 23 Octobre 2020 de 9h00 à 12h00

- Les résultats seront **encadrés**, et toute réponse devra être justifiée (sauf exception).
- Les exercices sont indépendants les uns des autres.
- Le barème est donné à titre indicatif, pour vous permettre de proportionner vos efforts. Il pourra être marginalement modifié. Noter que la somme excède 20.
- Le sujet comporte 4 pages.
- Le devoir dure 3h00.
- Les téléphones sont rangés éteints dans les sacs ; en guise d'aide mémoire, vous avez droit au recto d'une feuille manuscrite (par vos soins, pas de photocopie) sur la table.

Exercice 1. [Renversement du temps, 8 points]

Soit P une matrice de transition irréductible sur Ω, et π son unique mesure de probabilité stationnaire. On définit une matrice $\hat{P}$ par :

$$\forall x, y \in \Omega,\ \hat{P}(x,y) = \frac{\pi(y)P(y,x)}{\pi(x)}.$$

1. Montrer que $\hat{P}$ est bien définie et qu'il s'agit d'une matrice stochastique.
2. Que vaut $\hat{P}$ si P est réversible par rapport à π ? En déduire l'unique mesure de probabilité stationnaire de $\hat{P}$ dans ce cas. Vérifier que cette mesure de probabilité est encore l'unique mesure de probabilité stationnaire de $\hat{P}$ si l'on ne suppose pas la réversibilité de P.

Dans toute la suite on choisit $\Omega = \{0, \ldots, n\}$, et, avec la notation $x \wedge y = \min\{x, y\}$:

$$P(x,y) = \frac{1}{2}\left(\mathbb{1}_{\{y=(x+1)\wedge n\}} + \mathbb{1}_{\{y=0\}}\right),$$

3. Calculer l'unique mesure de probabilité stationnaire π de P.
4. Donner la matrice de transition $\hat{P}$. (On pourra faire un dessin des transitions possibles sur un graphe).
5. Montrer que $\hat{P}(0,\cdot) = \pi$.
6. Calculer $\hat{P}^t(x,\cdot)$ pour tout $0 \le t \le x \le n-1$.
7. Déduire sans calcul de ce qui précède $\hat{P}^t(x,\cdot)$ pour $t > x$, $0 \le x \le n-1$.
8. Soit $\hat{X}$ la chaîne de Markov de matrice de transition $\hat{P}$. On note

$$\tau_{n-1}(\hat{X}) = \min\{t \ge 0 : \hat{X}_t = n-1\}$$

le temps d'atteinte de $n-1$ par cette chaîne, calculer $\mathbb{P}_n(\tau_{n-1}(\hat{X}) = k)$ et reconnaître cette loi.

9. Soit $t \leq n$. En déduire $\hat{P}^t(n, n-1)$ puis plus généralement $\hat{P}^t(n, \cdot)$. Reconnaître en particulier la mesure de probabilité $\hat{P}^n(n, \cdot)$.
10. Déduire sans calcul de ce qui précède $\hat{P}^t(n, \cdot)$ pour tout $t > n$. Commenter ce résultat.

Correction. *P bien définie car l'irréductibilité implique que l'unique mesure stationnaire est strictement positive en tout point x ; le fait que P soit stochastique découle alors du fait que π est stationnaire pour P : ceci assure en effet que la somme sur les lignes vaut 1, (et les coefficients sont positifs ou nuls).*

Si P est réversible par rapport à P, alors $\pi(x)\hat{P}(x,y) = \pi(y)P(y,x)$ implique $\hat{P} = P$, dont π est l'unique mesure stationnaire. De façon plus générale, on vérifie sans peine que π est une mesure de proba stationnaire de $\hat{P}$, puisque :

$$\sum_x \pi(x)\hat{P}(x,y) = \sum_x \pi(x)\frac{\pi(y)P(y,x)}{\pi(x)} = \pi(y)\sum_x P(y,x) = \pi(y)$$

Ensuite, pour l'unicité de cette mesure stationnaire il suffit de noter que $\hat{P}$ est encore irréductible. Soit $x, y \in \Omega$. Alors il existe $t \in \mathbb{N}$ tel que $P^t(y,x) > 0$ de l'irréductibilité de P et ceci implique que $\hat{P}^t(x,y) > 0$, par exemple parce que $\hat{P}^t(x,y) = \frac{\pi(y)P^t(x,y)}{\pi(x)}$.

On commence maintenant la partie avec la forme explicite de P, et d'abord le calcul de π. On commence par observer que $\pi(x+1) = \pi(x)1/2$ pour tout $x \in \{0, \ldots, n-2\}$ et $\pi(n) = \pi(n-1)$. Ainsi $\pi(x) = (1/2)^{x \wedge (n-1)}\pi(0)$ pour tout $x \in \{0, \ldots, n\}$. Puis

$$1 = \sum_{0 \leq x \leq n} \pi(x) = \pi(0) \sum_{0 \leq x \leq n-1} (1/2)^{x \wedge n-1} = \pi(0)\left(\frac{1-(1/2)^n}{1-1/2} + (1/2)^{n-1}\right) = 2\pi(0).$$

Donc

$$\pi(x) = (1/2)^{(1+x)\wedge n}.$$

Ensuite, on calcule :

$$\hat{P}(x,y) = \mathbb{1}_{x=0}\pi(y) + \mathbb{1}_{0<x<n}\mathbb{1}_{y=x-1} + \mathbb{1}_{x=n}\frac{1}{2}(\mathbb{1}_{y=n-1} + \mathbb{1}_{y=n}).$$

On souligne que $\hat{P}(0, \cdot)$ est égal à π

On voit que $\hat{P}^t(x, \cdot) = \delta_{x-t}(\cdot)$ pour tout $0 \leq t \leq x \leq n-1$ (évolution déterministe). Aussi $\hat{P}^{x+1}(x, \cdot) = \pi$ puis comme π est encore une mesure stationnaire pour $\hat{P}$, $\hat{P}^t(x, \cdot) = \pi$ pour tout $t > x$. Pour comprendre $\hat{P}(n, \cdot)$, on note que $\mathbb{P}_n(\tau_{n-1} = k) = (1/2)^k$ pour tout $k \geq 1$, on reconnait une loi géométrique issue de 1 de paramètre de succès 1/2. Ainsi pour $x \geq 1$, $\hat{P}^t(n, n-x) = \mathbb{P}(\tau_{n-1} = t-(x-1)) = (1/2)^{t-x+1}$ d'où, pour $x \leq n-1$, $\hat{P}^n(n,x) = (1/2)^{n-(n-x-1)} = (1/2)^{x+1)}$, et donc puisque $\hat{P}^n$ est une mesure de proba qui coincide avec π sur $\{0, \ldots, n-1\}$, nécessairement, $\hat{P}^n(n, \cdot) = \pi$. Puisque π est stationnaire pour $\hat{P}$, on en tire $\hat{P}^x(n, \cdot) = \pi$ pour tout $x > n$. Ainsi la chaîne atteint sa mesure stationnaire en temps fini.

Exercice 2. [Marche aléatoire biaisée : temps d'atteinte, 10 points]

Soit deux réels $p, q \in]0,1[$ tels que $p + q = 1$, avec $p \neq q$. On considère sur $\Omega = \{0, 1, \ldots, n\}$ la marche aléatoire $(X_t)_{t \in \mathbb{N}}$ biaisée avec des conditions au bord de type réflexion, de matrice de transition :

$$\begin{cases} P(x,y) = p\mathbb{1}_{\{y=x+1\}} + q\mathbb{1}_{\{y=x-1\}} & \text{si } x \in \{1, \ldots, n-1\}, y \in \{0, \ldots, n\} \\ P(0,1) = P(n,n-1) = 1 \end{cases}$$

On rappelle que $\tau_x = \min\{t \geq 0 : X_t = x\} \in \mathbb{N} \cup \{\infty\}$ désigne le temps d'atteinte de x. On pose $f(x) = \mathbb{E}_x[\tau_n] \in [0, \infty]$ pour $x \in \{0, \ldots, n\}$.

1. Construire un réseau $\{G, c\}$ dont la marche aléatoire associée est la chaîne de Markov $(X_t)_{t \in \mathbb{N}}$
2. À l'aide de l'identité du temps de transport, calculer la valeur du temps de transport

$$t_{0 \leftrightarrow n} := \mathbb{E}_0[\tau_n] + \mathbb{E}_n[\tau_0].$$

On note que le temps de transport donne une borne supérieure sur $\mathbb{E}_0[\tau_n]$. Les questions qui suivent visent à établir une borne inférieure.

3. Justifier l'identité en loi suivante :

$$\tau_n \text{ sous } \mathbb{P}_0 \ = \sum_{i=1}^{G-1} T_i + R,$$

où
— G est une variable Géométrique (sur $\mathbb{N}^\star$) de paramètre $p = \mathbb{P}_0(\tau_0^+ > \tau_n)$
— $T_1, T_2, \ldots$ dont des variables i.i.d. de loi $\mathbb{P}_0(\tau_0^+ \in \cdot | \tau_0^+ < \tau_n)$
— R suit la loi $\mathbb{P}_0(\tau_n \in \cdot | \tau_0^+ > \tau_n)$,
ces variables étant toutes indépendantes. (Si $G = 1$, on convient que la première somme est nulle). On ne demande pas une démonstration rigoureuse. On sera concis et on pourra faire un dessin.

4. En déduire :

$$\mathbb{E}_0[\tau_n] = \left(\frac{1}{\mathbb{P}_0(\tau_0^+ > \tau_n)} - 1 \right) \mathbb{E}_0[\tau_0^+ | \tau_0^+ < \tau_n] + \mathbb{E}_0[\tau_n | \tau_0^+ > \tau_n]$$

5. Calculer $\mathbb{E}_0[\tau_0^+]$ et $\mathbb{P}_0(\tau_0^+ > \tau_n)$ à l'aide de relations vue en cours.
6. On admet que $\mathbb{E}_0[\tau_0^+ | \tau_0^+ < \tau_n] \to \mathbb{E}_0[\tau_0^+]$ quand $n \to \infty$. Déduire de ce qui précède un équivalent simple de $\mathbb{E}_0[\tau_n]$ dans le cas $p < q$ (cet équivalent ne nécessite pas de calculer $\mathbb{E}_0[\tau_n | \tau_0^+ > \tau_n]$).

La suite de l'exercice est indépendante de ce qui précède. Elle vise à calculer de façon exacte

$$f : \{0, \ldots, n\} \to [0, \infty], \ k \mapsto f(k) = \mathbb{E}_k[\tau_n]$$

à l'aide de la méthode à un pas.

7. Soit $k \in \{1, \ldots, n-1\}$. Donner un lien entre $\tau_n \circ \theta(X)$ et $\tau_n(X)$ sous $\mathbb{P}_k$ la loi de la chaîne issue de k[1], et en déduire à l'aide de la propriété de Markov une expression de $f(k)$ en fonction de $f(k-1)$ et $f(k+1)$.
8. Exprimer $f(0)$ en fonction de $f(1)$. Donner aussi $f(n)$.
9. On pose $\ell(k) = f(k+1) - f(k)$ pour $0 \leq k \leq n-1$. Montrer que $\ell(k)$ satisfait une équation de récurrence du type $\ell(k) = \alpha \ell(k-1) + \beta$, $1 \leq k \leq n-1$, pour des paramètres α et β que l'on précisera.

1. on rappelle que θ est l'opérateur de translation : $\theta(X)_t = X_{t+1}$, et plus généralement, $\theta_s(X)_t = X_{t+s}, t, s \in \mathbb{N}$

10. On rappelle que la solution de la récurrence affine précédente est de la forme

$$\ell(k) = \frac{\beta}{1-\alpha} + \alpha^k \left(\ell(0) - \frac{\beta}{1-\alpha}\right), \quad 0 \le k \le n-1.$$

Calculer la fonction ℓ dans notre cas et en déduire la fonction f.

11. Donner des équivalents de $\mathbb{E}_0[\tau_n]$ quand $n \to \infty$ dans les deux cas suivants :
(i) $p > q$
(ii) $p < q$
où p, q sont indépendants de n.

12. On pose, pour $\lambda \in \mathbb{R} \setminus \{0\}$ fixé, $q_n = 1 - p_n = \frac{1}{2} + \frac{\lambda}{n}$. Montrer qu'il existe alors $\varphi : \mathbb{R} \to \mathbb{R}^+$ que l'on précisera, tel que, quand $n \to \infty$:

$$\mathbb{E}_0[\tau_n] \sim n^2 \varphi(\lambda).$$

Correction. *Le réseau associé est donné par le graphe G spécifié avec les conductances :*

$$c(x, x+1) = (\frac{p}{q})^x$$

et $c(x,x) = 0$: pas de boucles. Il suffit en effet de vérifier que ce réseau donne les bonnes transitions : Pour $x \in \{1, \ldots, n-1\}$,

$$P(x, x+1) = \frac{(\frac{p}{q})^x}{(\frac{p}{q})^x + (\frac{p}{q})^{x-1}} = \frac{(\frac{p}{q})}{(\frac{p}{q}) + 1} = p$$

$$P(x, x-1) = \frac{(\frac{p}{q})^{x-1}}{(\frac{p}{q})^{x-1} + (\frac{p}{q})^x} = \frac{1}{1 + \frac{p}{q}} = q$$

tandis que $P(n, n-1)$ et $P(0,1)$ valent bien 1. On fera en sorte dans les écritures suivantes que les quantités qui interviennent soient positives quand $q > p$ et c'est dans ce régime qu'on prendra nos équivalents. Notons que $r(x, x+1) = (\frac{q}{p})^x$ et partant

$$\mathcal{R}(0 \leftrightarrow n) = \frac{(\frac{q}{p})^n - 1}{\frac{q}{p} - 1}, \quad \mathcal{C}(0 \leftrightarrow n) = \frac{\frac{q}{p} - 1}{(\frac{q}{p})^n - 1}$$

Il faut aussi calculer le coefficient

$$c_G = 2 \sum_x \sum_y c(x,y) = 2 \sum c(e) = 2 \frac{(\frac{p}{q})^n - 1}{\frac{p}{q} - 1}$$

On peut aussi désormais calculer le temps de transport

$$t_{0 \leftrightarrow n} = 2 \frac{(\frac{p}{q})^n - 1}{\frac{p}{q} - 1} \frac{1 - (\frac{q}{p})^n}{1 - \frac{q}{p}} \sim 2 \frac{pq}{(p-q)^2} (\frac{q}{p})^n \text{ si } q > p$$

Le tenps de transport donne une borne pour

$$\mathbb{E}_0[\tau_n] \le t_{0 \leftrightarrow n} \sim 2 \frac{pq}{(p-q)^2} (\frac{q}{p})^n$$

Il suffit pour justifier la décomposition demandée de noter que les différentes excursions de la trajectoire en dehors de 0 sont indépendantes : le nombre de telles excursions jusquà la première qui atteint n suit une loi Geométrique G de paramètre $p = \mathbb{P}_0(\tau_n < \tau_0^+)$*, et conditionnellement à cette valeur, G, les* $G-1$ *premières excursions sont indépendantes de loi celle de* τ_0^+ *sachant* $\tau_0^+ < \tau_n$ *et la G-ième excursion est de loi celle de* τ_0^+ *sachant* $\tau_0^+ > \tau_n$*.En prenant les espérances on a bien :*

$$\left(\frac{1}{\mathbb{P}_0(\tau_0^+ > \tau_n)} - 1\right)\mathbb{E}_0[\tau_0^+|\tau_0^+ < \tau_n] + \mathbb{E}_0[\tau_n|\tau_0^+ > \tau_n] = \mathbb{E}_0[\tau_n]$$

Ensuite :

$$\mathbb{P}_0(\tau_0^+ < \tau_n) = c(0)\mathbb{P}_0(\tau_0^+ < \tau_n) = \mathcal{C}(0 \leftrightarrow n) = \mathcal{C}(0 \leftrightarrow n) = \frac{\frac{q}{p} - 1}{(\frac{q}{p})^n - 1}$$

Aussi,

$$\mathbb{E}_0[\tau_0^+] = \frac{c_G}{c_0} = 2\frac{(\frac{p}{q})^n - 1}{\frac{p}{q} - 1}$$

donc en utilisant l'équivalent donné dans l'énoncé :

$$(\frac{1}{\mathbb{P}_0(\tau_0^+ > \tau_n)} - 1)\mathbb{E}_0[\tau_0^+|\tau_0^+ < \tau_n] \sim \frac{1}{\mathbb{P}_0(\tau_0^+ > \tau_n)}\mathbb{E}_0[\tau_0^+] = 2\frac{(\frac{q}{p})^n - 1}{\frac{q}{p} - 1}\frac{1 - (\frac{p}{q})^n}{1 - \frac{p}{q}} \sim \frac{2pq}{(q-p)^2}(\frac{q}{p})^n$$

soit le même équivalent que précédemment.

$\tau \circ \theta(X) = \theta(X) + 1$ *sur l'événement* $\{\tau \neq 0\}$*, où l'addition a lieu dans* $[0, +\infty]$*. Comme* $\mathbb{P}_k(\tau \neq 0) = 0$ *dès lors que* $k \notin \{0, n\}$*. Pour un tel k, on a alors que*

$$\begin{aligned} f(k) = \mathbb{E}_k[\tau] &= \mathbb{E}_k[\tau \circ \theta + 1] \\ &= \mathbb{E}_k[\tau \circ \theta + 1, X_1 = k+1] + \mathbb{E}_k[\tau \circ \theta + 1, X_1 = k-1] \\ &= \mathbb{E}_{k+1}[\tau + 1]\mathbb{P}_k(X_1 = k+1) + \mathbb{E}_{k-1}[\tau + 1]\mathbb{P}_k(X_1 = k-1) \\ &= 1 + p\mathbb{E}_{k+1}[\tau + 1] + q\mathbb{E}_{k-1}[\tau + 1] \\ &= 1 + pf(k+1) + qf(k-1) \end{aligned}$$

tandis que $f(n) = 0$*. Enfin, pour* $k = 0$ *on obtient :*

$$\begin{aligned} f(0) = \mathbb{E}_0[\tau] &= \mathbb{E}_0[\tau \circ \theta + 1] \\ &= \mathbb{E}_0[\tau \circ \theta + 1, X_1 = 1] \\ &= \mathbb{E}_1[\tau + 1]\mathbb{P}_0(X_1 = 1) \\ &= \mathbb{E}_1[\tau] + 1 \\ &= f(1) + 1 \end{aligned}$$

Ensuite, pour $k \notin \{0, n\}$*, si l'on pose* $\ell(k) = f(k+1) - f(k)$*, alors l'équation* $f(k) = 1 + pf(k+1) + qf(k-1)$ *se réécrit* $0 = 1 + p\ell(k) - q\ell(k-1)$ *soit encore :*

$$\ell(k) = \frac{q}{p}\ell(k-1) - \frac{1}{p}$$

Ceci est équivalent :

$$\ell(k) - \alpha = \frac{q}{p}(\ell(k-1) - \alpha)$$

avec $\alpha(1-\frac{q}{p}) = -\frac{1}{p}$, *c'est-à-dire :*

$$\alpha = \frac{-\frac{1}{p}}{1-\frac{q}{p}} = \frac{-1}{p-q} \text{ et } 1+\alpha = \frac{-2q}{p-q}$$

(On ne remplacera qu'à la fin α *par la valeur, c'est plus sûr). Notons aussi que* $\ell(0) = -1$*. Ainsi on obtient :*

$$\ell(k) = \alpha + (\frac{q}{p})^k(-1-\alpha)$$

Maintenant,

$$\begin{aligned} -f(k) = f(n) - f(k) &= \sum_{j=k}^{n-1} \ell(j) \\ &= \sum_{j=k}^{n-1} [\alpha + (\frac{q}{p})^k(-1-\alpha)] \\ &= (n-k)\alpha + (-1-\alpha)\frac{(\frac{q}{p})^k - (\frac{q}{p})^n}{1-\frac{q}{p}} \end{aligned}$$

c'est-à-dire :

$$f(k) = (n-k)(-\alpha) + (1+\alpha)\frac{(\frac{q}{p})^k - (\frac{q}{p})^n}{1-\frac{q}{p}}$$

Et il est commode d'écrire cette expression sous la forme :

$$f(k) = \begin{cases} \frac{1}{p-q}(n-k) - \frac{2pq}{(p-q)^2}\left[(\frac{q}{p})^k - (\frac{q}{p})^n\right] & \text{si } p > q \\ -\frac{1}{q-p}(n-k) + \frac{2pq}{(p-q)^2}\left[(\frac{q}{p})^n - (\frac{q}{p})^k\right] & \text{si } q > p \end{cases}$$

On en tire les équivalents pour p, q *fixés :*

$$f(0) \sim \begin{cases} \frac{n}{p-q} & \text{si } p > q \quad \text{(croissance linéaire)} \\ \frac{2pq}{(p-q)^2}(\frac{q}{p})^n & \text{si } q > p \quad \text{(croissance exponentielle)} \end{cases}$$

On en tire aussi les équivalents suivant pour $q_n = 1 - p_n = \frac{1}{2} + \frac{\lambda}{n}$ *:*

$$f(0) \sim \frac{n^2}{2}(\frac{e^{4\lambda}-1}{4\lambda^2} - \frac{1}{\lambda})$$

et on vérifie sans peine que la fonction de λ *dans le membre de droite est strictement positive en dehors de* 0.

Exercice 3. [Transformations de chaîne de Markov, 6+1 point]

Soit $(X_t)_{t\geq 0}$ une chaîne de Markov irréductible à valeurs dans Ω fini, de matrice de transition P de mesure stationnaire π. On pose

$$\begin{cases} \eta(X) = \min\{s \in \mathbb{N} : X_s \neq X_0\} \\ \eta_0(X) = 0, \quad \eta_{t+1}(X) = \eta_t(X) + \eta \circ \theta_{\eta_t}(X) \quad t \geq 0 \end{cases}$$

puis $Y_t = X_{\eta_t(X)}$ pour tout entier $t \in \mathbb{N}$.

1. Décrire brièvement en toute lettres comment la chaîne Y est construite à partir de X.
2. Calculer $\mathbb{P}_x(\eta \geq t)$ et en déduire $\mathbb{P}_x(\eta < \infty) = 1$. Observer que $(\mathbb{P}_x(X_\eta = y))_{x,y\in\Omega}$ est une matrice stochastique, en donner une expression à l'aide de P.
3. Montrer que, pour $y_0, y_1, \ldots, y_t \in \Omega$,

$$\mathbb{P}(Y_0 = y_0, \ldots, Y_{t-1} = y_{t-1}, Y_t = y_t) = \mathbb{P}(Y_0 = y_0, \ldots, Y_{t-1} = y_{t-1})Q(y_{t-1}, y_t)$$

   avec
$$Q(x,y) = \frac{P(x,y)}{1 - P(x,x)}\mathbb{1}_{x\neq y},$$

   et en déduire que $(Y_t)_{t\geq 0}$ est encore une chaîne de Markov.
4. Montrer que
$$\left(\frac{\pi(x)(1 - P(x,x))}{\sum_y \pi(y)(1 - P(y,y))}\right)_{x\in\Omega}$$

   est l'unique mesure de probabilité stationnaire de la chaîne $(Y_t)_{t\geq 0}$.

Soit de plus $A \subset \Omega$. On pose

$$\begin{cases} \tau^+(X) = \tau_A^+(X) = \min\{s \geq 1 : X_s \in A\} \\ \tau_0^+(X) = 0, \quad \tau_{t+1}^+(X) = \tau_t^+(X) + \tau^+ \circ \theta_{\tau_t^+}(X) \quad t \geq 0 \end{cases}$$

puis $Z_t = X_{\tau_t^+}$ pour tout entier $t \in \mathbb{N}$. (On pose $\tau^+ = \tau_A^+$ à la première ligne pour ne pas alourdir la notation sachant que A est fixé, et aussi pour éviter la confusion avec le cas d'un indice entier τ_t^+ ensuite défini).

5. Reprendre les questions 1 et 3, jusqu'à obtenir que $(Z_t)_{t\geq 0}$ est une chaîne de Markov pour une certaine matrice de transition R que l'on précisera.
6. (Bonus, plus difficile, à ne traiter que s'il vous reste du temps). Montrer que
$$\left(\frac{\pi(x)}{\pi(A)}\right)_{x\in A}$$

   est l'unique mesure de probabilité stationnaire de la chaîne $(Z_t)_{t\geq 0}$. On pourra commencer par montrer que, pour tout $t \geq 1, y \in A$:

$$\sum_{x\in A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ \leq t) + \sum_{x\notin A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) = \pi(y).$$

Correction. *La chaine Y est la chaîne trace qui consigne les différentes valeurs prises par la chaîne X en omettant les répétitions ; en conséquence, on s'attend à ce que la matrice de transition soit nulle sur la diagonale. Ensuite, $\mathbb{P}_x(\eta \geq t) = P(x,x)^{t-1}$ mais $P(x,x) < 1$ puisque la chaîne est irréductible, ceci entraine bien que $\mathbb{P}_x(\eta = +\infty) = \mathbb{P}_x(\eta \geq t) = P(x,x)^{t-1} \to 0$ donc $\mathbb{P}_x(\eta < +\infty) = 1 - \mathbb{P}_x(\eta = +\infty) = 1$. Ensuite, pour $y \neq x$*

$$\begin{aligned}
\mathbb{P}_x(X_\eta = y) &= \sum_{t\geq 1} \mathbb{P}_x(X_\eta = y, \eta = t) \\
&= \sum_{t\geq 1} \mathbb{P}_x(X_1 = x, \dots, X_{t-1} = x, X_t = y) \\
&= \sum_{t\geq 1} P(x,x)^{t-1} P(x,y) \\
&= \frac{P(x,y)}{1 - P(x,x)}
\end{aligned}$$

et cette matrice définit bien une matrice stochastique puisque $\sum_{y:y\neq x} P(x,y) = 1 - P(x,x)$ puisque P est elle-même stochastique. On peut noter que η_t est un temps d'arrêt en écrivant :

$$\{\eta_t \leq r\} = \{\sum_{s=1}^{r} \mathbb{1}_{X_{s-1}\neq X_s} \geq t\}$$

mais ceci n'est pas indispensable si on n'invoque pas la propriété de Markov forte (ce qu il est toujours possible de faire en décomposant selon les valeurs prises par le temps d'arrêt). On écrit donc :

$$\begin{aligned}
\mathbb{P}(Y_0 = y_0, \dots, Y_{t-1} = y_{t-1}, Y_t = y_t) &= \mathbb{P}(X_{\eta_0} = y_0, \dots, X_{\eta_{t-1}} = y_{t-1}, X_{\eta_{t-1}+\eta\circ\theta_{\eta_{t-1}}} = y_t) \\
&= \sum_r \mathbb{P}(X_{\eta_0} = y_0, \dots, X_r = y_{t-1}, \eta_{t-1} = r, X_{r+\eta\circ\theta_r} = y_t) \\
&= \sum_r \mathbb{P}(X_{\eta_0} = y_0, \dots, X_{\eta_{t-1}} = y_{t-1}, \eta_{t-1} = r) \mathbb{P}_{y_{t-1}}(X_\eta = y_t) \\
&= \mathbb{P}(X_{\eta_0} = y_0, \dots, X_{\eta_{t-1}} = y_{t-1}) \mathbb{P}_{y_{t-1}}(X_\eta = y_t) \\
&= \mathbb{P}(X_{\eta_0} = y_0, \dots, X_{\eta_{t-1}} = y_{t-1}) Q(y_{t-1}, y_t)
\end{aligned}$$

en notant à la troisièmne ligne que : $\{X_{\eta_0} = y_0, \dots, X_{\eta_{t-1}} = y_{t-1}, \eta_{t-1} = r\} \in \sigma\{X_0, \dots, X_r\}$ ce qui permet d'appliquer Markov. et on en tire par une récurrence finie immédiate que

$$\mathbb{P}(Y_0 = y_0, \dots, Y_{t-1} = y_{t-1}, Y_t = y_t) = \mathbb{P}(X_0 = x_0) Q(y_0, y_1) \dots Q(y_{t-1}, y_t)$$

c'est-à-dire que $(Y_t)_t$ est une chaîne de Markov de matrice de transition Q. Ensuite, pour y fixé :

$$\begin{aligned}
\sum_x \pi(x)(1 - P(x,x)) R(x,y) &= \sum_x \pi(x)(1 - P(x,x)) \frac{P(x,y)}{1 - P(x,x)} \mathbb{1}_{x\neq y} \\
&= \sum_x \pi(x) P(x,y) \mathbb{1}_{x\neq y} \\
&= \pi(y) - \pi(y)P(y,y) = \pi(y)(1 - P(y,y))
\end{aligned}$$

d'où la conclusion en renormalisant par $\sum_z \pi(z)(1 - P(z,z))$.

Ensuite, on fait de même pour la suite (Z_t) de variables aléatoires, qui correspond aux valeurs que la chaîne restreinte à A. On note que τ_A^+ est un temps d'arrêt fini p.s. par irréductibilité de la chaîne (c'est le cours qui le garantit, on a même que ces temps sont d'espérance finie), on a de même e que τ_t^+ est encore un temps d'arrêt fini p.s. :

$$\begin{aligned}
\mathbb{P}(Z_0 = z_0, \ldots, Z_{t-1} = z_{t-1}, Z_t = z_t) &= \mathbb{P}(X_{\tau_0^+} = y_0, \ldots, X_{\tau_{t-1}^+} = z_{t-1}, X_{\tau_t^+} = z_t) \\
&= \mathbb{P}(X_{\tau_0^+} = y_0, \ldots, X_{\tau_{t-1}^+} = z_{t-1}, X_{\tau_{t-1}^+ + \tau \circ \theta_{\tau_{t-1}^+}} = z_t) \\
&= \sum_r \mathbb{P}(X_{\tau_0^+} = y_0, \ldots, X_r = z_{t-1}, \tau_{t-1}^+ = r, X_{\tau_{t-1}^+ + \tau^+ \circ \theta_{\tau_{t-1}^+}} = z_t) \\
&= \sum_r \mathbb{P}(X_{\tau_0^+} = y_0, \ldots, X_r = z_{t-1}, \tau_{t-1}^+ = r)\mathbb{P}_{z_{t-1}}(X_{\tau^+} = z_t) \\
&= \mathbb{P}(Z_0 = z_0, \ldots, Z_{t-1} = z_{t-1})\mathbb{P}_{z_{t-1}}(X_{\tau^+} = z_t)
\end{aligned}$$

et donc on pose $R(x,y) = \mathbb{P}_x(X_{\tau^+} = y)$ matrice stochastique sur A, et on en tire par une récurrence finie immédiate que :

$$\mathbb{P}(Z_0 = z_0, \ldots, Z_{t-1} = z_{t-1}, Z_t = z_t) = \mathbb{P}(X_0 = x_0)R(y_0, y_1)\ldots R(y_{t-1}, y_t)$$

Pour la mesure stationnaire (BONUS) c'est un peu plus compliqué : on commence montrer la formule proposée par récurrence sur t : on pose donc l'hypothèse de récurrence

$$H_t : \text{"} \sum_{x \in A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ \leq t) = + \sum_{x \notin A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) = \pi(y)\text{"}.$$

H_1 vaut :

$$\sum_{x \in A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ \leq 1) + \sum_{x \notin A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = 1) = \sum_{x \in A} \pi(x)P(x,y) + \sum_{x \notin A} \pi(x)P(x,y) = \pi(y)$$

Le calcul suivant montre que si H_t vaut, alors H_{t+1} vaut également :

$$\begin{aligned}
&\sum_{x \notin A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) \\
&= \sum_{x \notin A} \Big(\sum_{x_{-1} \in A} \pi(x_{-1})P(x_{-1}, x) \Big) \mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) \\
&= \sum_{x_{-1} \notin A} \sum_{x \notin A} \pi(x_{-1})P(x_{-1}, x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) + \sum_{x_{-1} \in A} \sum_{x \notin A} \pi(x_{-1})P(x_{-1}, x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) \\
&= \sum_{x_{-1} \notin A} \pi(x_{-1})\mathbb{P}_{x_{-1}}(X_{\tau_A^+} = y, \tau_A^+ \geq t+1) + \sum_{x_{-1} \in A} \pi(x_{-1})\mathbb{P}_{x-1}(X_{\tau_A^+} = y, \tau_A^+ = t+1)
\end{aligned}$$

H_t étant prouvé, on passe à la limite en t dans cette égalité : pour cela, on se rappelle que l'événement $\{\tau_A^+ < \infty\}$ est $\mathbb{P}_x$ p.s. pour tout $x \in A$, et donc :

$$\sum_{x \in A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ \leq t) \to \sum_{x \in A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ < +\infty) = \sum_{x \in A} \pi(x)\mathbb{P}_x(X_{\tau_A^+} = y)$$

mais aussi

$$\sum_{x \notin A} \pi(x) \mathbb{P}_x(X_{\tau_A^+} = y, \tau_A^+ = t) = \max_{x \notin A} \mathbb{P}_x(\tau_A^+ \geq t) \to 0.$$

Il suit :

$$\sum_{x \in A} \pi(x) \mathbb{P}_x(X_{\tau_A^+} = y) = \pi(y)$$

et il suffit alors de renormaliser par $\pi(A)$ *pour avoir une mesure de proba.*