Temps de mélange pour les transpositions adjacentes

L'objectif de ce texte est d'étudier le temps de mélange d'une chaîne de Markov dont l'espace des états est l'ensemble $\mathfrak{S}(N)$ des permutations de taille $N \geq 2$. La chaîne en question correspond à l'expérience aléatoire suivante. On considère une rangée de livres dans une bibliothèque : les livres sont numérotés de 1 à N, et initialement ordonnés de gauche à droite dans l'ordre $123\ldots N$. Le lecteur vient régulièrement choisir deux livres adjacents de cette bibliothèque, en positions i et $i+1$ (avec $i \in [\![1, N]\!]$) ; il les replace ensuite dans la bibliothèque en échangeant leurs deux positions, c'est-à-dire en replaçant le livre qui était en position i à la position $i+1$, et *vice et versa*. On convient que si $i = N$, alors le lecteur ne vient prendre que le livre en position N, et il le replace au même endroit.

On suppose que les choix des entiers $i = i_n$ sont indépendants à chaque étape $n \in \mathbb{N}_*$, et uniformes dans $[\![1, N]\!]$. On note $\sigma_n(k)$ le numéro du livre placé en position k après n étapes ; chaque fonction $\sigma_n : [\![1, N]\!] \to [\![1, N]\!]$ est une bijection dans $\mathfrak{S}(N)$, et on a par hypothèse $\sigma_0 = \mathrm{id}_{[\![1,N]\!]}$. La suite de permutations aléatoires $(\sigma_n)_{n\in\mathbb{N}}$ est une chaîne de Markov, et on va montrer que sa loi converge vers la mesure uniforme sur les permutations, en un temps au moins de l'ordre de $O(N^3 \log N)$.

1 Propriétés de la chaîne de Markov

Pour $i \in [\![1, N-1]\!]$, on note τ_i la transposition $(i, i+1)$:

$$\tau_i(j) = \begin{cases} i+1 & \text{si } j = i, \\ i & \text{si } j = i+1, \\ j & \text{sinon.} \end{cases}$$

1. Montrer que $(\sigma_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur $\mathfrak{S}(N)$, dont la matrice de transition est :

$$P(\sigma, \sigma') = \begin{cases} \frac{1}{N} & \text{si } \sigma' = \sigma, \\ \frac{1}{N} & \text{si } \sigma' = \sigma \circ \tau_i \text{ pour un } i \in [\![1, N-1]\!], \\ 0 & \text{sinon.} \end{cases}$$

2. On rappelle que le groupe symétrique $\mathfrak{S}(N)$ est engendré par les transpositions (i, j) avec $1 \leq i < j \leq N$: (i, j) échange i et j et laisse tous les autres entiers $k \in [\![1, N]\!] \setminus \{i, j\}$ invariants. Pour tout couple (i, j) avec $i < j$, donner une expression de cette transposition comme produit de transpositions adjacentes τ_a.

3. Montrer que la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$ est irréductible sur $\mathfrak{S}(N)$ et apériodique. Montrer aussi que la matrice P est bistochastique. En déduire que la mesure invariante de la chaîne est la mesure uniforme donnée par
$$\pi(\sigma) = \frac{1}{N!}$$
pour toute permutation $\sigma \in \mathfrak{S}(N)$.

4. On pose $\pi_n(\sigma) = \mathbb{P}[\sigma_n = \sigma]$. Donner pour tout $\sigma \in \mathfrak{S}(N)$ la valeur de $\lim_{n\to\infty} \pi_n(\sigma)$.

5. Programmer avec `Python/SageMath` la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$. On pourra stocker les permutations σ_n sous forme de listes `Python` $[\sigma_n(1), \ldots, \sigma_n(N)]$, ou utiliser la classe `Permutation` de `SageMath`. Illustrer le résultat de la question précédente, par exemple avec des permutations de taille $n = 5$.

2 La méthode de Wilson

La suite du problème est consacrée à l'étude de la vitesse de convergence des lois π_n vers leur limite. Si μ et ν sont deux probabilités sur un ensemble mesurable $(\mathfrak{X}, \mathscr{F})$, on rappelle que leur distance en variation totale est

$$\|\mu - \nu\| = \sup_{A\in\mathscr{F}} |\mu(A) - \nu(A)|.$$

6. Soit μ et ν deux mesures de probabilité sur un ensemble mesurable $\mathfrak{X}$, et $f : \mathfrak{X} \to \mathfrak{Y}$ une application. On note $f_*\mu$ et $f_*\nu$ les mesures images de μ et ν par f. Montrer que

$$\|f_*\mu - f_*\nu\| \leq \|\mu - \nu\|.$$

7. Soit $f : \mathfrak{X} \to \mathbb{R}$ une application mesurable telle que

$$\int_{\mathfrak{X}} (f(x))^2\, \mu(dx) < +\infty \qquad ; \qquad \int_{\mathfrak{X}} (f(x))^2\, \nu(dx) < +\infty.$$

Ceci permet de considérer l'espérance et la variance de f sous la loi μ et sous la loi ν. On note

$$m = \min(\mathbb{E}_\mu(f), \mathbb{E}_\nu(f));$$
$$\sigma^2 = \max(\mathrm{Var}_\mu(f), \mathrm{Var}_\nu(f)),$$

et on suppose que

$$|\mathbb{E}_\mu[f] - \mathbb{E}_\nu[f]| \geq r\,\sigma$$

pour une certaine constante $r > 0$. Utiliser la question précédente pour montrer que sous cette hypothèse,

$$\|\mu - \nu\| \geq 1 - \frac{8}{r^2}.$$

Étant donnée la partie mesurable $(m + \frac{r\sigma}{2}, +\infty) = B \subset \mathbb{R}$, on pourra utiliser l'inégalité de Chebyshev pour contrôler $(f_*\mu)(B)$ et $(f_*\nu)(B)$.

Soit P une matrice stochastique irréductible apériodique sur un ensemble fini $\mathfrak{X}$, et π la mesure invariante de P, qu'on suppose réversible. On rappelle que dans ce cas les valeurs propres de P sont des nombres réels dans l'intervalle $(-1, 1]$. On considère une fonction non nulle $f : \mathfrak{X} \to \mathbb{R}$ que l'on représente par un vecteur colonne $(f(x))_{x\in\mathfrak{X}}$, et qui vérifie les deux hypothèses suivantes :

(a) La fonction f est un vecteur propre à droite pour P et pour une valeur propre λ avec $\frac{1}{2} < \lambda < 1$:

$$Pf = \lambda\, f.$$

(b) Pour tout $x \in \mathfrak{X}$, la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ issue de x et de matrice de transition P vérifie

$$\mathbb{E}_x[(f(X_1) - f(x))^2] \leq R$$

pour une certaine constante R.

On fixe $x \in \mathfrak{X}$ point de départ de la chaîne $(X_n)_{n\in\mathbb{N}}$, et on note $D_n = f(X_{n+1}) - f(X_n)$.

8. Montrer que $\mathbb{E}_x[D_n \,|\, X_n = y] = (\lambda - 1)\, f(y)$ et que $\mathbb{E}_x[(D_n)^2 \,|\, X_n = y] \leq R$. En déduire que

$$\mathbb{E}_x[(f(X_{n+1}))^2] \leq R + (2\lambda - 1)\, \mathbb{E}_x[(f(X_n))^2].$$

9. Montrer que, pour tout entier $n \in \mathbb{N}$,

$$\mathrm{Var}_x(f(X_n)) \leq \frac{R}{2(1-\lambda)}.$$

Montrer que $\mathrm{Var}_\pi(f)$ vérifie la même inégalité.

10. On note π_n la loi de X_n sous $\mathbb{P}_x$. Calculer $\mathbb{E}_{\pi_n}[f]$ et $\mathbb{E}_\pi[f]$. En utilisant les résultats des questions précédentes, montrer que

$$\|\pi_n - \pi\| \geq 1 - \frac{4R}{(1-\lambda)\lambda^{2n}\,|f(x)|^2}.$$

3 Spectre de la chaîne des transpositions adjacentes

On considère de nouveau la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$ de la première partie du problème.

11. Fixons $a \in [\![1, N]\!]$. Montrer que $((\sigma_n)^{-1}(a))_{n\in\mathbb{N}}$ est une chaîne de Markov sur $[\![1, N]\!]$ de matrice de transition

$$Q(k,l) = \begin{cases} \frac{1}{N} & \text{si } l = k-1, \\ \frac{1}{N} & \text{si } l = k+1, \\ 1 - \frac{1}{N} & \text{si } l = k \in \{1, N\}, \\ 1 - \frac{2}{N} & \text{si } l = k \in [\![2, N-1]\!]. \end{cases}$$

Que représente concrètement la valeur $\sigma_n^{-1}(a)$?

12. Montrer que $f(k) = \cos(\frac{(2k-1)\pi}{2N})$ est une fonction propre de Q, de valeur propre

$$\lambda = 1 + \frac{2}{N}\left(\cos\frac{\pi}{N} - 1\right).$$

Montrer que pour $N \geq 4$, cette valeur propre vérifie

$$1 - \frac{\pi^2}{N^3} < \lambda < 1 - \frac{\pi^2}{N^3}\left(1 - \frac{1}{N^2}\right)$$

et satisfait aux hypothèses de la méthode de Wilson.

13. Soit $F : \mathfrak{S}(N) \to \mathbb{R}$ la fonction définie par :

$$F(\sigma) = \sum_{a=1}^{N} f(a)\, f(\sigma^{-1}(a)).$$

Montrer que F est un vecteur propre de la matrice P pour la valeur propre λ de la question précédente. Montrer qu'on a aussi

$$F(\mathrm{id}_{[1,N]}) = \frac{N}{2}.$$

14. Montrer que pour toute permutation σ,

$$\mathbb{E}_\sigma[(F(\sigma_1) - F(\sigma))^2] \leq \frac{4\pi^2}{N^2}.$$

On pourra remarquer que f est $\frac{\pi}{N}$-Lipschitzienne, et que si $(\sigma_n)_{n\in\mathbb{N}}$ est la chaîne de Markov issue de $\sigma_0 = \sigma$, alors dans $F(\sigma_1)$, au plus deux termes diffèrent de ceux de la somme $F(\sigma)$.

15. En appliquant l'inégalité de la question 10. avec $N \geq 4$, montrer que

$$\|\pi_n - \pi\| \geq 1 - \frac{KN}{(\lambda)^{2n}}.$$

pour une certaine constante $K > 0$. En déduire que pour tout $\varepsilon > 0$, il existe une constante réelle C_ε telle que

$$(\|\pi_n - \pi\| \leq \varepsilon) \quad \Rightarrow \quad \left(n \geq \frac{N^3(\log N + C_\varepsilon)}{2\pi^2}\right).$$

Ainsi, il faut un temps au moins égal à $\frac{N^3 \log N(1+o(1))}{2\pi^2}$ pour mélanger correctement les livres. Il a été démontré en 2013 par Lacoin que le résultat ci-dessus est précis, au sens suivant : si $n \geq (1+\eta)\,\frac{N^3 \log N}{2\pi^2}$ pour $\eta > 0$, alors réciproquement $\|\pi_n - \pi\|$ est petit.